Univerzita Karlova

Fakulta sociálních věd

Institut ekonomických studií

Jana Votápková

Essays on Health Economics and Health Policy

Disertační práce

Praha 2020

Charles University in Prague

Faculty of Social Sciences

Institute of Economic Studies

Jana Votápková

Essays on Health Economics and Health Policy

Dissertation thesis

Prague 2020

Author: PhDr. Jana Votápková
Advisor: doc. PhDr. Martin Gregor, Ph.D.
Academic year: 2019/2020

Abstract

This dissertation thesis is a collection of three applied economics essays and a comprehensive discussion essay which includes policy recommendations resulting from the empirical essays included in this dissertation thesis, and other early work of the author. The topics analyzed were motivated by recent Czech healthcare reforms that address the recommendations raised by the OECD. Firstly, efficiency of Czech general hospitals is analyzed stressing the need for a well-functioning DRG system. Secondly, inpatient user charges which were introduced in 2008 and abolished shortly after, are analyzed as a tool which increases private participation in healthcare expenses and thus may boost healthcare efficiency both through additional resources and reduction of excess demand for healthcare services. Thirdly, personal savings, pre–retirement savings in particular, are viewed as another private source which could ease the tight public healthcare and pension budgets. It is shown that health status is a positive determinant for pre–retirement liquid savings, thus preventive health programs should be given a priority within the Czech healthcare system.

Keywords:

Healthcare, inpatient sector, efficiency, user charges, pre–retirement savings.

Abstrakt

Tato disertační práce je souborem tří esejí zabývajících se ekonomií zdravotnictví v České republice. Součástí je také esej, která shrnuje jejich výsledky do doporučení pro zdravotní politiku České republiky. Výběr témat obsažených v této disertační práci byl motivován doporučeními OECD v oblasti zdravotnictví. První esej analyzuje efektivitu všeobecných nemocnic v ČR a zdůrazňuje nezbytnost dobře fungujícího systému DRG. Poplatky za pobyt v nemocnici, tj. za lůžkovou péči, které byly zavedeny v roce 2008 a krátce poté byly opět zrušeny, jsou analyzovány jako nástroj, který zvyšuje spoluúčast pacientů a může tedy účinně zvýšit efektivitu zdravotního systému prostřednictvím dodatečných zdrojů financí a redukcí nadměrné poptávky po zdravotních službách. Osobní úspory, zvláště osob v před-důchodovém věku, jsou v disertaci chápány jako další soukromý zdroj vstupující do veřejného systému, který může ulevit napjatému rozpočtu zdravotnictví a důchodového systému. Bylo zjištěno, že zdravotní stav obyvatel v před-důchodovém věku má pozitivní vliv na výši likvidních úspor tohoto segmentu obyvatel.

Klíčová slova:

Zdravotnictví, lůžková péče, efektivita, poplatky ve zdravotnictví, před-důchodový věk.

I hereby declare that this dissertation thesis is my original work and that all sources and literature used are listed.

Signature

Acknowledgment

Herewith, I would like to express my deepest gratitude to doc. PhDr. Martin Gregor, Ph.D., my thesis supervisor, for his time, inevitable remarks and advice he has given me throughout the process of writing of this thesis and his support he has provided me with during my entire Ph.D. studies. His support and patience were the primary factors that determined the completion of this thesis.
I would like to thank to all my co-authors, PhDr. Lenka Šťastná, Ph.D., Mgr. Pavlína Žílová and Camilla Mastromarco, Ph.D. for their comments and inputs to our work.
I also greatly benefited from numerous discussions with my colleagues from the Institute of Economic Studies, as well as, the participants at various seminars, conferences and summer schools.
I very much appreciated comments and suggestions of the reviewers to the pre-defense version of the thesis - Olena Stavrunova, Ph.D., Andrea Menclová, Ph.D. and Michal Horný, Ph.D., MSc.
Financial support of the grants received throughout my studies is highly appreciated, namely, GAUK 910 915 (Chapters 1 and 2), GAUK 336 215 (Chapter 3), Grant Agency of the Czech Republic No. P402/12/G097, TAČR Omega TD03000209 and financial support of the European Commission (Chapter 4).

Contents

List of Figures

List of Tables

List of Abbreviations

CZ	Czech Republic
CZSO	Czech statistical office
ČSSD	Social democratic party of the Czech Republic
DiD	Difference in differences
DMU	Decision-making unit
DRG	Diagnostic related group
EU	European Union
EU-15	Members of EU prior 2004
EU-28	All EU members as of 2018
EU-13	Countries accessing after 2004
EU-SILC	Survey of income and living conditions
FCS	Fully conditional specification
FDH	Free disposable hull
GDP	Gross domestic product
GP	General practitioner
HTA	Health technology assessment
LSCV	Least squares cross-validation
MDC	Major diagnostic category
NHS	National health system
NRC	Národní referenční centrum
OECD	Organization for Economic Cooperation and Development
PCA	Principal component analysis
PPS	Purchasing power standard
RAND HIE	RAND Health insurance experiment
SFA	Stochastic frontier analysis
SHARE	Survey of Health, Aging and Retirement in Europe
UZIS	Institute for Health Information and Statistics, CZ
WHO	World Health Organization
2SLS	Two-Stage Least Squares

Introduction

The Czech healthcare system is a statutory health insurance system with obligatory wage-based contributions. Private participation is minimal. The system has been undergoing moderate reforms, including corporatization of hospitals, user-charge introduction, and electronization of the system. Given population aging, the system is expected to be under increasing fiscal pressure overtime.

This dissertation thesis consists of four essays, three of which empirically analyze the segments of the Czech healthcare system which have been reformed lately or which still wait to be improved. The last essay discusses the results of this dissertation thesis and provides a comprehensive policy recommendations arising from my research both within and beyond this dissertation thesis. The thesis is directly motivated by the recent health policy reform attempts of the Czech government, which responded to the recommendations of the OECD (OECD, 2011, 2014). Results of the research lead to policy implications that broadly correspond with the recent policy recommendations of the OECD (2018) for the healthcare system of the Czech Republic.

The OECD (2018) stresses that the Czech Republic should be more financially sustainable. In order to reach this ultimate goal, the OECD (2018) recommends that the Czech Republic should (1) deal with the challenge of aging population, (2) improve hospital management, (3) make availability and quality of care more equal across regions, (4) reimburse hospitals based on performance rather based on a flat rate and (5) strengthen the role of prevention and the role of General Practitioners in the care for the chronically ill.

The financial sustainability recommendation penetrates into the implications of my entire research within and beyond this thesis. I suggest that more efficient hospital management and effective use of the DRG will save resources on inpatient care (Chapter 1). If private participation in healthcare expenses increases, both through individual savings and cost-sharing, the public budget should be more financially sustainable, the more so in the view of population aging (Chapater 2 and Chapter 3). Chapter 3 deals with the pre-retirement segment of population exclusively and points to the important role of prevention finding out that health status is a positive determinant of pre-retirement savings. Thus, consistent with the OECD (2018) recommendation I argue that prevention should be given a priority within the Czech healthcare system.

The thesis is structured into four chapters three of which correspond to the empirical essays. Chapter 4 discusses and connects the results of this dissertation research and includes policy recommendations resulting from my work both within and beyond this thesis. The references from all chapters are grouped into a single list at the end of the dissertation thesis. Responses to reviewers' suggestions and an overview of incorporated changes to the pre-defense version of the thesis are included in the Appendix.

Most of the essays have already been published, with the exception of Chapter 2. In addition, the policy impact of my research has been recognized by OECD (2018) who cited my previous work (Votápková & Žílová, 2012, 2016b; Alexa *et al.*, 2015) which is beyond the scope of this dissertation thesis.

Chapter 1 analyzes efficiency of Czech general hospitals applying the conditional efficiency approach. The essay was published as Mastromarco, C., Šťastná, L & Votápková, J. (2019) Efficiency of hospitals in the Czech Republic: conditional efficiency approach, *Journal of Productivity Analysis*, 51, 73-79 (The article was edited by William H. Greene, NYU). It is a follow-up research to my master thesis which was published as Votápková, J. & Šťastná, L. (2013) Efficiency of hospitals in the Czech Republic, *Prague Economic Papers*, 4.

The sample of hospitals, the time period analyzed and variables used in these two studies overlap only partly. Votápková & Šťastná (2013) use a pooled panel of 99 hospitals observed between 2001–2008 and Mastromarco *et al.* (2019) analyze 81 hospitals over the period 2006–2010. The total of 78 hospitals overlap between studies. Mastromarco *et al.* (2019) use an improved list of variables (DRG and publication outputs) which were found desirable but missing in the former.

Consistent with OECD (2018) it was found that the DRG, even though not yet optimally designed, very much improves benchmarking of hospitals. The DRG is still only a performance measure, not yet the reimbursement mechanism in the Czech Republic. Historical prices (flat rate) are still used as the reimbursement mechanism instead. I believe that implementation of the DRG into the reimbursement system for inpatient care will very much improve efficiency of this sector.

The essay is co-authored by Camilla Mastromarco (5%) who revised the methodological section for the final submission to the *Journal of Productivity Analysis*; and Lenka Šťastná (10%) who revised the text throughout the process of writing and provided comments on my intended further steps from the beginning to the end. The essay is identical with the published paper.

Chapter 2 analyzes the effect of inpatient user charges on the amount of inpatient care provided. The essay exploits the dataset of Mastromarco *et al.* (2019) and Votápková & Šťastná (2013) analyzing inpatient user charges from the supply side perspective. It is assumed that hospitals respond to fixed demand from the community rather than they endogeneously determining the demand. The introduction of user charges in 2008 is in line with the financial sustainability recommendation which has been long stated by the

OECD. However, after vivid political discussions, regions governed by Social Democracy started to reimburse patients for user charges in hospitals with their major political influence in 2009. These included regional joint-stock companies and some municipal nonprofit and for-profit hospitals. Such a set-up enables a natural experiment. In the essay, I apply the difference-in-differences approach testing whether after the 'abolition' of user charges for inpatient care in regional hospitals, the number of inpatient days in these hospitals increased. Only a small and statistically unstable effect was found, consistent with the findings in other Czech research analyzing the effect of user charges, although mostly in different healthcare sectors (Kalousová, 2014; Votápková & Žílová, 2016a; Hromádková, 2016; Zápal, 2010).

Chapter 2 is a sole work of mine. The essay will shortly be submitted to the IES Working Papers series. Chapter 2 is a follow-up research of the effect of user charges which started with the analysis of outpatient care in Votápková & Žílová (2016a) which was published in *Health Economics Review*.

Chapter 3 investigates whether health status is a determinant of pre-retirement liquid savings. Wave 5 of the SHARE dataset, in particular its pre-retirement segment, is analyzed. The essay addresses the aging of population and financial sustainability challenges raised by OECD (2018). By 2050, the share of 65+ is expected to rise from 18 % to 30 %. Healthcare expenses devoted to this segment of population are expected to rise from 45 % to 75 % (OECD, 2018). In order for the Czech healthcare system to be financially sustainable in the future, private participation in healthcare expenses should most probably increase since current private participation in healthcare expenses in the Czech Republic is among the lowest. Large-scale public financing of healthcare expenses may otherwise put the public healthcare (but also pension) system at risk very soon. The elderly will thus have to secure themselves for their expenses for retirement.

We used a 2SLS methodology with liquid savings being the exogenous variable and health status being endogenous. A Wu Hausman test was applied to test endogeneity within the model which proved 2SLS as appropriate as opposed to the OLS. In addition, we failed to reject endogeneity of health status with respect to saving and exogeneity of savings with respect to health status also by extracting residuals from each reduced form equation and plugging them into the other structural equation.

We found that health status determines the level of pre-retirement liquid savings. The government should thus give priority to preventive health programs. Not only is prevention often cheaper than treatment per se, but well–targeted prevention decreases the number of acute admissions, decreases the length of stay, decreases rate of re–hospitalization and thus improves efficiency of hospitals and the overall healthcare system. In addition, a higher savings rate has macroeconomic implications for investments and growth.

The essay was published as Votápková, J. & Žílová, P. (2018) Health status as a determinant for pre-retirement liquid savings, IES Working Papers 10/2018. No major changes have occurred since then. It is co-authored by Pavlína Žílová whose contribution was approx. 10 %. She participated in variable definition and data extraction. She also revised and edited the text.

Chapter 4 is a comprehensive discussion essay which provides policy recommendations. It summarizes policy implications of the research within this dissertation thesis and also our closely related research not covered in the thesis. Most of the recommendations are in line with OECD (2018). Having introduced stylized facts that put the Czech health sector into the international context, we summarize implications for the Czech inpatient sector. Section 4.1, compares two sets of results of our research, specifically Votápková & Šťastná (2013) and Mastromarco *et al.* (2019). The comparison suggests that the management and reimbursement structure of Czech hospitals should be improved. We showed that the DRG system is extremely beneficial for hospital benchmarking. In Section 4.1, potential benefits of the new CZ-DRG system as a reimbursement mechanism are highlighted. Once the DRG enters the reimbursement system, the current flat rate reimbursement should be abandoned and efficiency of the inpatient sector is expected to increase. We argue that the DRG reimbursement system will also improve efficiency of capital investments of hospitals. Section 4.2 summarizes implications resulting from our research dealing with increased private participation on healthcare expenses in Chapters 2 and 3 and Votápková & Žílová (2016a). Section 4.2 also highlights the need for a strengthened role of prevention resulting from our research in Chapter 3.

Chapter 4 is based on Votápková, J. & Žílová, P. (2015) Healthcare efficiency in the Czech Republic - evidence for inpatient care *IES Occasional Papers* no. 2 but was significantly modified. Stylized bacts were shortened and moved to the Appendix of Chapter 4, Section 4.1 was edited because the paper Mastromarco *et al.* (2019) was accepted into the *Journal of Productivity Analysis* only in January 2019 being under revisions there from 2014. As the conditional efficiency analysis changed since Votápková & Žílová (2015), so did the discussion in Section 4.1. Section 4.1 was edited so that it includes the latest information about the DRG which has changed since 2015. Additional recommendations resulting from our research carried out since 2015 were added. Section on efficiency of capital investment was incorporated into Section 4.1 and section analyzing short-term hospitalizations and day cases was excluded from the dissertation thesis.

The IES OP 2/2015 was co-authored by Pavlína Žílová, whose contribution to it was approx 15 %. Her main role was in expert methodological consultations and text revisions and provided the view of the Ministry of Health, particularly in Section 4.1. She also provided access to some unique sets of data.

Chapter 1

Efficiency of hospitals in the Czech Republic: Conditional efficiency approach

Inpatient care consumes the majority of hospital resources (Yong & Harris, 1999). For Czech hospitals, inpatient costs represent around 50% of total costs on average. Outpatient care accounts for 15–20% of total costs, the rest is taken up by transportation costs and non-medical expenses.

The aim of this essay is to evaluate the efficiency of hospitals in the Czech Republic during 2006–2010, i.e. how efficiently they transform inputs into outputs, and explain differences in efficiency based on several hospitals' characteristics. Measuring the efficiency of hospitals has become widespread within individual countries in the last decades. Recent evidence is available from Portugal (Ferreira & Marques, 2016; Ferreira *et al.*, 2018), Germany (Tiemann & Schreyögg, 2012), Greece (Halkos & Tzeremes, 2011), the Netherlands (Blank & Valdmanis, 2010) and Nordic countries (Linna *et al.*, 2010). Varabyova *et al.* (2017) compare efficiency of hospitals in Italy and Germany. Efficiency is assessed also for hospitals in the U.S. (Bates *et al.*, 2006; Clement *et al.*, 2008; Nayar & Ozcan, 2008), Sweden (Janlöv, 2007), Switzerland (Farsi & Filippini, 2006), Austria (Hofmarcher *et al.*, 2002), or Great Britain (Jacobs, 2001). More examples can be found in a recent overview of studies by Kohl *et al.* (2018) and in earlier works by Hollingsworth (2008) or Worthington (2004).

The efficiency of hospitals—the way they transform inputs into outputs—may be affected by environmental factors, which are beyond the scope of hospital management. Operating in a good/bad environment increases/decreases a hospital's efficiency. Hence, environmental factors should be taken into account in the efficiency estimation (Blank & Valdmanis, 2010).

In non-parametric estimations, there are several ways to account for environmental variables (see e.g. Fried *et al.*, 2008). As is shown in Simar & Wilson (2007), traditional two-stage approaches suffer from several problems (see for example Matranga *et al.* (2014), Tiemann & Schreyögg (2012) for applications of two-stage approaches in health care). Even though Simar & Wilson (2007) propose a method to overcome complications of two-stage approaches (for example Araújo *et al.* (2014) analyzing Brazilian hospitals follow their approach), we recognize merits of a one-stage conditional efficiency approach (originally developed by Cazals *et al.* (2002), extended by Daraio & Simar (2005, 2007)) most suitable to account for environmental variables in the efficiency analysis of Czech hospitals. We follow the conditional efficiency model that allows us to distinguish between continuous and discrete environmental variables and, at the same time, does not require separability between the environmental and input-output spaces.

In the sphere of healthcare, analyses applying the conditional-efficiency approach are rare. Halkos & Tzeremes (2011) apply conditional efficiency to healthcare provision in Greek regions, Cordero *et al.* (2015) analyze primary care providers using a conditional approach, Varabyova *et al.* (2017) use a conditional approach to compare efficiency in Italian and German hospitals. Hence, this paper is among a few healthcare studies applying a conditional efficiency model. In addition, the paper extends previous research on Czech hospitals—non-parametric analyses in Dlouhý *et al.* (2007) and Novosádová & Dlouhý (2007) who did not account for environments at all, and a parametric analysis in Votápková & Šťastná (2013)—by using the most appropriate non-parametric method and by covering more recent and better data on outputs not available before (Diagnostic-Related-Groups, DRG, reflecting the severity of treated patients, which is currently being developed in the Czech Republic).

In this essay, we focus on inpatient care and evaluate how the total inpatient costs are transformed into outputs which include the total number of patients treated at acute wards weighted by the DRG case-mix index, patients treated at nursing wards, and the number of nurses per one bed which represents a qualitative indicator of treatment. On top of that, we include the number of weighted publications similar to Linna & Häkkinen (1998) and Linna (1998) to reflect not only research production, but also involvement in teaching, especially when university hospitals with higher inpatient costs support research activities.

We aim to explain differences in efficiency using a set of environmental variables, such as nonprofit ownership, presence of highly specialized centers, teaching status (university hospital), occupancy rate or specific time effects. A non-parametric significance test and partial regression plots are employed to uncover the significance and the direction of the effect of these variables. We find that nonprofit hospitals, university hospitals, and hospitals with specialized centers have generally lower efficiency. Additionally, efficiency worsens in the years 2009–2010 since additional revenues received in the form of user charges directly from patients allowed hospitals to spend more and loosen their budget constraints—the effect was strong particularly for nonprofit hospitals consistent with

Newhouse (1970).

This essay is organized as follows. Section 1.1 provides the theoretical background for conditional efficiency analysis and describes the methodology of the non-parametric significance test and partial regression plots. Section 1.2 presents the dataset and introduces the variables employed. Section 1.3 presents the results and Section 1.4 concludes.

1.1 Methodology

Consider a production technology with a set of all feasible inputs $x \in \mathbb{R}^p_+$ and outputs $y \in \mathbb{R}^q_+$ and denote $Z \in \mathbb{R}^d_+$ several environmental factors exogenous to the production process itself, but which may explain a part of it. Following Cazals *et al.* (2002) and Daraio & Simar (2005), the unconditional (marginal) attainable set of feasible combinations of inputs and outputs, $\Psi = \{(x,y) \in \mathbb{R}^{p+q}_+ | x \text{ can produce } y\}$ can be characterized by $\Psi = \{(x,y) | H_{X,Y}(x,y) > 0\}$, where $H_{X,Y}(x,y) = \text{Prob}(X \leq x, Y \geq y)$. So Ψ is the support of the joint random variable (X,Y). The best practice frontier follows from Ψ, which is freely disposable, but unknown in reality, and has to be estimated from a random sample. Let $N = (1,\ldots,n)$ be the set of decision-making units (DMUs) in the dataset. We analyze the problem from the input-oriented perspective, because hospital management has greater control over inputs than outputs.

From the frontier literature it is known that, assuming free disposability of inputs and outputs, the unconditional (marginal) input oriented Farrell-Debreu technical efficiency of a production plan (x,y), may be defined as:

$$\theta(x,y) = \inf\{\theta | (\theta x, y) \in \Psi\} = \inf\{\theta | S_{X|Y}(\theta x|y) > 0\}, \tag{1.1}$$

where $S_{X|Y}(x|y) = \text{Prob}(X \leq x | Y \geq y)$ is the nonstandard conditional survival function of X given that $Y \geq y$.

For conditional efficiency measures we define the attainable set $\Psi^z \subset \mathbb{R}^{p+q}_+$ as the support of the conditional probability:

$$H_{X,Y|Z}(x,y|z) = \text{Prob}\left(X \leq x, Y \geq y \,|\, Z = z\right).$$

Accordingly, the conditional input oriented technical efficiency of a production plan $(x,y) \in \Psi^z$, facing conditions z, is defined in Daraio & Simar (2005) as:

$$\theta\left(x,y|z\right) \;=\; \inf\left\{\theta | (\theta x, y) \in \Psi^z\right\} = \inf\{\theta | S_{X|Y,Z}(\theta x|y,z) > 0\}, \tag{1.2}$$

where $S_{X|Y,Z}(x|y,z) = \text{Prob}(X \leq x | Y \geq y, Z = z)$.

Nonparametric estimators of the attainable sets can be obtained by plugging nonparametric estimators of the survivor functions in the definitions above. Plugging the empirical version of $S_{X|Y,Z}$ in (1.2) provides the popular FDH (Free Disposal Hull) estimator of Ψ. A nonparametric estimator of the conditional survival function $S_{X|Y,Z}(x|y,z)$ could be obtained by using standard smoothing methods where a bandwidth h has to be determined for each component of (Z).

Daraio & Simar (2005, 2007) and Bădin *et al.* (2010) discuss in detail how to choose the appropriate bandwidths. They are determined by the estimation of conditional distributions $S_{X|Y,Z}(x|y,z)$, where we condition on $Y \geq y$ and a particular value of $Z = z$ and here standard tools from Hall *et al.* (2004) and Li & Racine (2008) can be adapted. Of course here only the variables (z) require smoothing and appropriate bandwidths, since we have:

$$\widehat{S}_{X|Y,Z}(x|y,z) = \frac{\sum_{i=1}^{n} \mathbb{1}(x_i \leq x, y_i \geq y) K_{h_z}(z_i)}{\sum_{i=1}^{n} \mathbb{1}(y_i \geq y) K_{h_z}(z_i)}, \tag{1.3}$$

where the functions $K_{h_z}(z_i)$ are kernels (see Bădin *et al.* (2010) for technical details).[1]

In order to estimate the conditional distributions in (1.3), we redefine the components of the multivariate Z to include discrete variables (Bădin *et al.*, 2010, 2012), such that $z_i = (z_i^c, z_i^u)$, $i = 1, \ldots, n$, where $z_i^c \in \mathbb{R}^v$ is a vector of continuous environmental variables, and $z_i^u \in \mathbb{R}^w$ is a vector of unordered discrete variables.[2] The generalized product kernel function is obtained as a multiplication of the standard multivariate product kernel functions of each of the groups of variables, such that:

$$K_{h_z}(z_i) = \prod_{s=1}^{r} K_{h_s^c}^c(z_s^c - z_{is}^c) \prod_{s=r+1}^{r+w} K_{h_s^u}^u(z_s^u, z_{is}^u) \tag{1.4}$$

where $K_{h_s^c}^c(\cdot)$ and $K_{h_s^u}^u(\cdot)$ are univariate kernel functions and h_s^c and h_s^u are bandwidths for continuous and unordered discrete environmental variables, respectively. For continuous variables, we use Epanechnikov kernel which has a compact support, i.e. $K_{h_s^c}^c(z_s^c - z_{is}^c) = 0$ if $|z| \geq 1$ and Aitchison & Aitken (1976) is used for discrete univariate kernel functions for unordered discrete variables. As a method of bandwidth selection for both continuous and discrete variables, we apply the least squares cross-validation method (Bădin *et al.*, 2010, 2012) based on the closely related conditional probability density functions as suggested by Li & Racine (2008) and developed by Hall *et al.* (2004).[3]

These nonparametric estimators are consistent with rate $n^{1/(p+1)}$ and Weibull limiting distribution for the unconditional FDH (see Park *et al.*, 2000). For the conditional case, we have similar results where n is replaced by nh^d where d is the dimension of all the conditioning variables (Z), so $d = r + p + 2$ (see Jeong *et al.*, 2010). So the rates of convergence of the conditional estimators are deteriorated by the dimension d.

In applied studies, the application of these nonparametric techniques may be problematic because the presence of outliers or extreme data points in real data samples,

[1] Optimal bandwidths can be selected by Least Squares Cross-Validation (LSCV) or by Maximum Likelihood Cross-Validation, which are asymptotically equivalent, see e.g. Li & Racine (2007).

[2] The model may be extended to also include ordered discrete variables (DeWitte & Kortelainen, 2013).

[3] The bandwidth refers to vector of bandwidths containing individual bandwidths for each variable.

which fully determine the estimated frontier and the measurement of inefficiencies, are totally unrealistic. To solve this problem, approaches have been proposed in the frontier literature (Cazals *et al.*, 2002; Daouia & Simar, 2007) to keep all the observations in the sample but to replace the frontier of the empirical distribution by (conditional) quantiles or by the expectation of the minimum (or maximum) of a subsample of the data. This latter method defines the order-m frontier that we will use here.

To be short, the partial output-frontier of order-m is defined for any integer m and for an output y, as the expected value of the minimum of the input of m units drawn at random from the populations of firms producing more output than y. Formally:

$$\theta_m(x,y) = \mathbb{E}\left[\min_{1,\dots,m}\left\{\max_{j=1,\dots,p}\left(\frac{X_i^j}{x^j}\right)\right\}\right], \tag{1.5}$$

where the X^j are independently distributed as $S_{X|Y}(\cdot|Y \geq y)$. The same applies for the conditional order-m frontier where the X^j are distributed as $S_{X|Y,Z}(\cdot|Y \geq y, Z = z)$. Nonparametric estimators are obtained by plugging the nonparametric estimators of the survival functions in (1.5).

Cazals *et al.* (2002) shows that, when m increases and converges to ∞, the order-m frontier and its estimator converge to the full frontier. For a finite m, the frontier will not envelop all the data points and so is much more robust than the FDH to outliers and extreme data points.[4] Another advantage of these estimators is that they achieve the parametric rate of convergence $\sqrt{n}$ and that they have a normal limiting distribution.

1.1.1 Impact of environmental variables on the production process

To find out the influence of environmental variables on the production process, we follow the procedure described in Bădin *et al.* (2012) and we will compare estimates of $\theta_m(x,y|z)$ with those of $\theta_m(x,y)$, i.e. using 'conditional' and 'unconditional' efficiency scores. The procedure allows disentangling the potential effects of environmental variables on the boundary (shift of the frontier) and on the distribution of the inefficiencies (see Bădin *et al.*, 2012; Mastromarco & Simar, 2015). The first effect can be investigated by considering the ratios of conditional to unconditional efficiency measures, which are measures relative to the full frontier of the conditional and the unconditional attainable sets respectively. As illustrated in Daraio and Simar (2007), some extreme or outlying data points may hide the real effect of Z, so it is suggested to do the analysis with order-m frontier, with large values of m to get robust estimates of the full frontier. In this case, the ratios to be analyzed are given by:

$$\hat{R_i^z} = \frac{\hat{\theta}_m(x_i, y_i|z_i)}{\hat{\theta}_m(x_i, y_i)} \tag{1.6}$$

[4] Daouia & Gijbels (2011) analyze these estimators from a theory of robustness perspective.

As stated in Bădin *et al.* (2012), the full frontier ratios, or their robust version with large values of m, indicate only the influence of Z on the shape of the frontier, whereas the partial frontiers for small values of m, characterizes the behavior of the shift more in the center of the distribution of efficiencies, inside the attainable sets.[5] A tendency of the ratios to decrease with the conditioning variables indicates a favorable effect of these variables on the distribution of the efficiencies and the opposite in the case of an unfavorable effect.[6]

We have a sample of n pairs $(z_i, \hat{R}(x_i, y_i|z_i)), i = 1, \dots n$, in the nonparametric model, we estimate the local average of $\hat{R}(x_i, y_i|z_i)$, the localization of which is determined by the bandwidth h (see Bădin *et al.*, 2014). Following Racine & Li (2004) and Daraio & Simar (2014) we use kernel weighted local linear least squares, a non-parametric regression technique which smoothes both continuous and discrete variables without sample splitting.[7]

The following local linear least squares minimization problem has to be solved:

$$\min_{\hat{\alpha},\hat{\beta}} \sum_{i=1}^{n} (\hat{R}_i^z - \hat{\alpha}(z_i) - \hat{\beta}(z_i))^2 K_{h_z}(z_i), \tag{1.7}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are local linear estimators to be obtained, such that $\hat{\alpha} = \hat{\alpha}(z)$ and $\hat{\beta} = \hat{\beta}(z)$ are consistent estimators of the true conditional mean function $f(z) = E(Q^z|z)$ and the gradient $\beta(z) = \frac{\partial E(Q^z|z)}{\partial z}$. Additionally, $K_{h_z}(\cdot)$ is the generalized product kernel function as in (1.4) and h_z is the bandwidth vector again estimated by the least-squares cross-validation method (Li & Racine, 2004).

Then we test significance of each continuous and discrete variable (Racine, 1997; Racine *et al.*, 2006).

1.1.2 Partial Regression Plots

We follow Daraio & Simar (2005, 2007) and visualize the effects of Z in partial regression plots. In our multivariate setting, we plot $\hat{R}_i^z$ against one variable fixing all other variables (at the median).

The interpretation of the regression line (in case of input orientation) is the following:

(i) If the regression line is increasing, vector Z is detrimental (unfavorable) to efficiency. According to Daraio & Simar (2005), the environmental variable here acts like 'extra' undesired output requiring more inputs in the production activity, hence Z

[5]For instance, if $m = 1$, the order-m frontier turns out to be an average production function and the ratios (1.6) would analyze the shift of the mean of the distribution of the inefficiencies.

[6]As explained in Bădin *et al.* (2012), the ratios are not bounded by 1, because the order-m efficiency scores are not bounded by 1.

[7]Note that Li *et al.* (2016) propose a complete smoothing technique which allows for different bandwidth parameters for continuous variables in different categories of the discrete variables.

exerts a negative effect on the production process. Unconditional efficiency is lower for larger values of Z—hence, $\hat{R}_i^z$ will increase on average with Z.

(ii) If the regression line is decreasing, then Z is conducive (favorable) to efficiency. Here, the environmental variable works as a 'substitutive' input to the production process, allowing the DMU to save inputs in the production process, i.e. environmental factors inherently reduce the amount of inputs hospitals require to treat their patient base. Unconditional efficiency is greater for larger values of Z—hence, R_i^z will decrease when Z increases.

1.2 Data

Data on 81 general hospitals for the period 2006–2010 was analyzed. From the total number of Czech general hospitals, 36% were excluded for various reasons: some of the hospitals were closed during the period examined, incorporated into larger entities, or did not report data.[8] Outlier-detection analysis as of Wilson (1993) and careful visual inspection of the data excluded additional 17 observations. An observation was not excluded if it was an outlier with extreme data, note in Figure 1.1 that extreme observations were left in the sample. However, if there was an obvious mistake stemming from the recording process, i.e. an observation did not fit into the time series of the hospital, it was excluded, leaving an unbalanced panel. The final unbalanced panel consists of 389 observations. The number of observations in each cross-section varies from 75 in 2007 and 2008 to 81 in 2010. Most of the hospitals treat up to 20,000 patients a year on average. There are two very big hospitals in the sample treating more than 70,000 patients a year. The third biggest hospital treats only 59,000 patients a year. The distribution of hospitals in terms of average size is depicted in Figure 1.1.

Data on individual hospitals was obtained from multiple sources,[9] data expressed in monetary terms, i.e. costs and salaries, were adjusted for inflation using an annual growth rate of inflation with base year 2006. Results were estimated with R 2.14.0 (R Development Core Team, 2006).

1.2.1 Input and output variables

The analysis focuses on the cost efficiency of inpatient care in hospitals. The only input variable in the analysis is total operating costs (*costs*) which comprise all inpatient costs

[8] Each record missing in the main data source was consulted with annual reports of the hospital. However, without information on costs or/and the number of DRG adjusted patients, the observation could not by definition be included.

[9] Most of the data was obtained from the Institute of Health Information and Statistics of the Czech Republic ('UZIS')[10]; Narodni referencni centrum ('NRC') provided us with data on Diagnostic-Related Groups ('DRG'); the Web of Science was used to retrieve data on publications affiliated to the particular hospital. Data on environmental characteristics was obtained from the Czech Statistical Office, Registry of Companies of the Czech Republic, and the Ministry of Health.

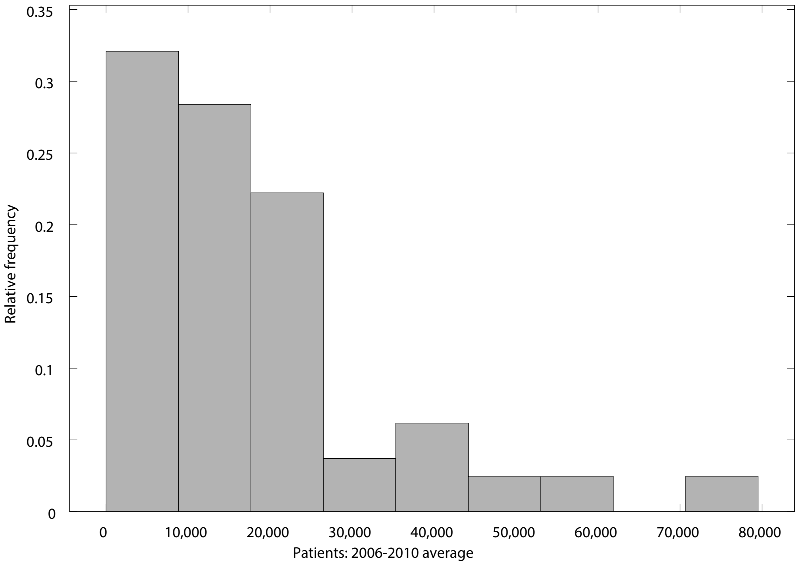


Figure 1.1. Distribution of hospitals by size

excluding capital costs. It was calculated as the multiplication of operating costs per inpatient day, the number of admissions, and the average length of stay (all publicly available from UZIS). UZIS calculates operating costs per inpatient day, C, as:

$$C = C_{In} \times \frac{1 + \frac{C_T + C_O + C_N}{C_{In} + C_{Out}}}{D},$$

where C_{In} are costs for inpatient care, C_T costs for medical transport, C_O costs for other medical care, C_N costs for non-medical procedures, C_{Out} outpatient costs and D number of inpatient days.

Hospitals produce an increase to patient's health status, however such output is technically impossible to capture. Hospital efficiency studies thus use intermediate outputs instead. The most important output for hospital efficiency analysis is the number of patients (only inpatients in our case, not outpatients) which is often used in the literature and which is preferred to inpatient days that may bias the results due to possible endogeneity born by the length of stay (see Zuckerman *et al.*, 1994; Farsi & Filippini, 2006; Hofmarcher *et al.*, 2002).[11]

Prior to the analysis, we first divided the number of patients into acute care and nursing care, because costs on acute care and nursing care significantly differ.

Cases within nursing care are rather homogeneous, but within acute care some hospitalizations are more expensive than others. Newhouse (1970) compares it to an aggregation problem which is inherent to a multiple-product firm. Not accounting for

[11] Chirikos & Sear (2000) additionally distinguish the first day of hospitalization from the rest assuming that most resources are consumed on the first day of hospitalization.

heterogeneity among acute-care hospitalizations may lead to bias in efficiency measurement (Ferreira & Marques, 2016; Bruning & Register, 1989; Burgess & Wilson, 1995). Weighting outputs according to case-mix has been acknowledged as vital, particularly when the sample consists of hospitals of different sizes, or university hospitals together with other acute hospitals, to minimize intra-hospital as well as inter-hospital differences (Chowdhury *et al.*, 2014; Rosko & Chilingerian, 1999; Valdmanis, 1992; Hofmarcher *et al.*, 2002).

Case-mix adjustment has to capture essential structural differences between hospitals (Anthun *et al.*, 2017). Different case-mix criteria appear in the literature, such as diagnostic-related groupings (Chowdhury & Zelenyuk, 2016; Hofmarcher *et al.*, 2002; Vitaliano & Toren, 1996; Magnussen, 1996); service-mix index (Ferreira & Marques, 2016) which is similar to the case-mix index but easier to compute; the types of patients treated (Kooreman, 1994), or country-specific weights (Chowdhury *et al.*, 2014).

Often case-mix is used as a weighting device, but sometimes case-mix is used as a separate output (Grosskopf & Valdmanis, 1993; Kooreman, 1994; Rosko & Chilingerian, 1999). Note that the choice of weighting criteria influences the distribution of efficiency scores (Anthun *et al.*, 2017; Magnussen, 1996; Chowdhury *et al.*, 2014) or may shift the frontier per se (Ferreira & Marques, 2016). Different methods in case-mix adjustment were tested on Portuguese hospitals in Ferreira & Marques (2016), including case-mix adjustment and service-mix adjustment. Interestingly, Ferreira & Marques (2016) find that proper environmental correction may strongly substitute for case-mix adjustment, but obviously at higher costs for the researcher. Case-mix adjustment of patients proved vital even when the efficiency of a specific treatment (lung cancer) was compared among hospitals in Beck *et al.* (2018).

Case-mix is used as a weighting device of acute inpatient care in this essay. We end up with two outputs related directly to inpatient care: (i) the number of acute care patients weighted for the DRG case-mix index (*acute_DRG*), and (ii) the number of patients in nursing care (*nursing*).

Quality of hospital output is important too. By definition of the efficiency model, quality usually represents another dimension of outputs, thus increasing efficiency of those hospitals which produce less output of higher quality, i.e. units which would be considered as inefficient without this additional output dimension. Quality of care has been accounted for differently in the literature. For instance, Zuckerman *et al.* (1994) used mortality rates, Vitaliano & Toren (1996) applied technology index and occupancy rate, Frohloff (2007) employed nurse-bed ratio.

The more nurses attend one bed per day, the higher the quality of care is expected. Indeed, Aiken *et al.* (2002) and Needleman *et al.* (2011) give evidence that a high patient-nurse ratio is associated with decreased health outcomes. The number of nurses per available bed (*nurse_bed*) thus represents a qualitative indicator which is used as a separate output in is essay, similar to Frohloff (2007). We are aware of the potential bias when the number of nurses is excessively large, driving the value-added on quality of the additional nurse to zero. This is not the case of Czech hospitals, which more often face shortages of nurses.

University hospitals incur additional costs for inpatient care because of teaching and research. Not only is the presence of students costly, but university hospitals are usually pioneers of new, but expensive technologies, to be able to teach their students the latest progress in medicine. Oftentimes, there are professors who, besides working as doctors, teach and are involved in research.

Data on the number of students/graduates affiliated with a particular university hospital could reflect the demanding nature of teaching, but unfortunately it is not available. Hence, we focus on research activity and include a variable accounting for publications by a hospital. Assuming that primarily big and university hospitals carry out research, publication output should improve low relative efficiency scores of a group of big and university hospitals found in Votápková & Šťastná (2013).

There has been a wide discussion in the literature whether to include teaching and research variables among outputs or among environmental variables (e.g. Vitaliano & Toren, 1996; Rosko & Chilingerian, 1999; Rosko, 2001, etc.). Since Czech hospitals themselves decide how much they will be involved in research activities, publications will be included among outputs similar to Linna (1998) and Linna *et al.* (1998) to reflect another kind of output which a hospital can control and which cannot be captured by the volume and case-mix variables (Vitaliano & Toren, 1996).

The fourth output variable (*publish*) is obtained as the first principal component of the data retrieved from the Web of Science database where inputs to the principal component analysis[12] are (i) articles, (ii) meeting abstracts, (iii) letters, reviews, proceedings papers, all weighted by the share of domestic authors affiliated to the particular hospital.[13]

We applied the positive affine transformation to avoid negative values and the minimum was added to all observations in the sample. The first principal component explains 64.45% of information in the publication data, while weights assigned to journal articles and monographs are almost the same. Table 1.1 presents the outcome of the principal

[12]Principal component analysis (PCA) transforms a large set of variables into a lower set of linearly uncorrelated values called principal components which best explain the variance in the data (Pearson, 1901).

[13]We performed the analysis also for different specifications of publication output. We first considered only journal articles from the Web of Science database, however some hospitals were found to produce more proceedings papers and their publication output would then be undervalued. In addition, we took into account publications from Czech research and innovations databases, however data is available only for university hospitals and hospitals receiving a grant from the Czech Ministry of Education.

component analysis.

Table 1.1. Principal component analysis

	PC1	PC2	PC3
Eigenvalue	1.934	0.994	0.072
Proportion	0.645	0.331	0.024
Cumulative	0.645	0.976	1.000
Components' loadings			
Journal articles	0.706	0.021	0.708
Monographs	0.698	0.149	−0.700
Other publications	0.120	−0.989	−0.091

Notes: Components' loadings of each variable explains the correlation between the particular component and the variable. Squared loading of each variable is then the proportion of the variance of the variable explained by the particular component.

1.2.2 Environmental characteristics

The environment in which hospitals operate may influence their efficiency. Hospitals may be managed differently when they are joint-stock companies instead of nonprofit institutions; university hospitals provide a different structure of services; hospitals with highly specialized treatment may incur higher costs in general.

In 2004, the process of corporatization of Czech hospitals started, the main purpose of which was to allocate resources more efficiently. Many hospitals were transformed from nonprofit institutions into joint-stock companies. The decision to corporatize was purely political, thus external to the hospital management decisions or hospital characteristics.

Until 2000, most of Czech hospitals were state-owned. After Czech regions ('kraje') were established in January 2000, many of the state-owned hospitals were transferred to their ownership. Some hospitals were overtaken by municipalities.'Kraje' and municipalities then decided to corporatize the hospitals into joint-stock companies hoping in more transparent accounting and financial stability of the hospitals. If the administrative unit decided to corporatize, it did so for all the hospitals in their hands. However, not all regions or municipalities decided to corporatize.

Even corporatized hospitals, however, are effectively under the public control since regions, district or municipalities are their major shareholders. Having carefully examined individual hospitals, it has been found that only 10% of hospitals in the sample are for-profit entities in private hands serving only 5 % of all patients. Expressed as a share of private entities, private hospitals in private hand make 23 % serving only 13 % of patients. Majority of private hospitals are thus corporatized entities which serve disproportionately more patients than their counterparts in parivate hands. Therefore, we consider only the nonprofit status (*nonprofit*) using a dummy variable taking the value of 1 when a hospital is public nonprofit and 0 otherwise, disregarding the different typess of private ownership.

We expect Czech nonprofit hospitals to be less efficient than the for-profit ones since the objective of for-profit hospitals in the Czech Republic is (i) to control financial flows and not to create losses due to stricter budget constraints (nevertheless, for-profit hospitals with a public authority as a major shareholder may ask for easing their debt under certain circumstances). Another objective of for-profit hospitals should be to (ii) increase technical efficiency. Thus, wards with low occupancy are often closed down. On the contrary, smaller public nonprofit hospitals more often keep wards that are not fully used to guarantee access of care as the objective of nonprofit hospitals is accessibility and quality output maximization (Newhouse, 1970; Pauly & Redisch, 1973; Feldstein, 1971; Lee, 1971; Reder, 1965) rather then profit and efficiency maximization.

In the Czech environment, specialized and teaching hospitals belong to nonprofit hospitals to guarantee them some sort of assurance of stability, but specialization and teaching variables are controlled for in the analysis.

We include a dummy variable for the presence of a specialized center (*specialization*) in a hospital, as of a list obtained from the Czech Ministry of Health.[14] Highly specialized treatment may be on the one hand connected with increased costs (not entirely captured by DRG adjusted output), which would decrease relative efficiency. On the other hand, doctors involved in specialized treatment may have higher publication activity, which would increase relative efficiency. The effect of this variable on efficiency will depend on which of these two directions dominates. There are 26 hospitals (corresponding to 114 observations in the pooled panel) with a specialized center in our sample.

Besides publishing results of research (captured in outputs), university hospitals reveal a different structure of services providing less basic and more highly-specialized care, management, and organization of resources (Vitaliano & Toren, 1996). Costs of university hospitals are often higher than costs in other hospitals (Grosskopf *et al.*, 2001b). University hospitals also suffer from congestion, i.e. excess use by residents. Grosskopf *et al.* (2001b) discovered that about 20 % of inefficiency of university hospitals is caused by congestion. Grosskopf *et al.* (2001a) carries out a comparison of the technical efficiency of university and non-university hospitals finding that only about 10 % of university hospitals can effectively compete with other hospitals based on provision of patients' services.

University status (*university*) included among determinants in this essay captures how the 'historic mission' affects a hospital's position vis-a-vis the best practice production frontier. The status of university hospital is assumed to exert a negative pressures on efficiency.

Occupancy rate (*occupancy*) defined as the ratio of the actual inpatient days to the maximum inpatient days possible, captures whether the hospital operates below its po-

[14] Examples include oncology centres, rheumatology centres, ophthalmology centres, etc.

tential capacity. Higher occupancy rate is expected to exert a positive effect on efficiency because hospitals face fixed costs connected with each bed available.

Out of the covered period 2006–2010, the efficiency of the last two years 2009 and 2010 could be influenced by two factors. The more important is the legislative change which came into force in 2008 introducing user charges for each inpatient day in a hospital and for outpatient visits, both regular and emergency.[15] Higher revenues soften budget constraints for a hospital, which may then afford higher operating costs. In such a case, we would expect a decrease of efficiency in these two years.

On the contrary, fiscal stress that spread due to the world financial crisis is assumed to work mostly in the opposite direction. Hospitals as well as other public and private institutions are forced to save money, hence their costs should be lower (efficiency for given outputs should increase). These two contradictory effects may also balance out resulting in no special effect upon efficiency.

We include a dummy variable taking the value of 1 for 2009 or 2010, and zero otherwise (*2009_2010*). The effect of the dummy will show whether hospitals were affected by the fiscal crisis, or whether user charges made up for the shortage of finances.[16]

Additionally, we test whether nonprofit and university hospitals behave differently in years 2009 and 2010. In two robustness checks we include the interaction of the dummy for 2009_2010 and nonprofit status (*2009_2010* × *nonprofit*); and 2009_2010 and university status (*2009_2010* × *university*). Descriptive statistics of all variables is provided in Table 1.2.

Table 1.2. Descriptive statistics of inputs, outputs and invironmental charactesistics

	Mean	Median	Min	Max	St.Dev.
Costs (thousands CZK)	629,000	338,000	61,900	3,840,000	788,000
Acute_DRG	21,755.10	12,426.28	1730.26	126,906.80	25,155.90
Nursing	249.867	175.356	0	1177.914	276.308
Nurse_bed	0.522	0.504	0.271	1.291	0.112
Publish	0.480	0	0	9.878	1.384
Nonprofit	0.545	1	0	1	0.498
Specialization	0.293	0	0	1	0.455
University	0.141	0	0	1	0.348
Occupancy	0.713	0.709	0.495	0.897	0.076
2009_2010	0.411	0	0	1	0.492
2009_2010 × University	0.057	0	0	1	0.231
2009_2010 × Nonprofit	0.221	0	0	1	0.415

[15] We assume that the effect of 2008 user charges is delayed to 2009, consistent with Zápal (2010)
[16] Individual year dummies were excluded based on preliminary analyses.

1.3 Empirical results

In this section, we present and discuss the empirical results of the analysis of Czech hospitals. We perform unconditional and conditional order-m analyses. In the conditional analysis, we account for environmental characteristics in which hospitals operate, hence compared to the unconditional efficiency score, the conditional efficiency score of a particular hospital is lower/higher if the hospital operates in favorable/detrimental environments. Therefore, conditional analysis reveals whether an environmental characteristic has a positive or a negative effect on efficiency.

We analyze the pooled dataset, i.e. a single frontier is constructed and hospitals are simultaneously compared among one another and across time. To check the poolability of the panel (to test whether the frontier is stable over time), we carry out preliminary unconditional efficiency analyses for each year and for a pooled dataset, and compute the Spearman's rank correlation coefficient between single year scores and the scores from the pooled dataset. Correlations vary from 0.87 in 2007 to 0.73 in 2010 and reveal a considerable time stability except for the years 2009 and 2010 (coefficients 0.76 and 0.73, respectively), which will be accounted for in the analysis.

Every non-parametric efficiency analysis is highly sensitive to outliers. Holding $m = 100$ to obtain the order-m scores, each observation out of 389 is compared to a random set of 100 observations.[17] The excessively large efficiency value above 1 would suggest that an observation lies far above the frontier, hence may be an outlier. As Table 1.4 shows, we do not detect any significant outliers in the sample (the maximum efficiency score is 1.32 and 1.04 in unconditional and conditional analysis, respectively).

Firstly, we present the results of the conditional analysis controlling for several environmental variables which may be beyond the scope of hospital management. Hence, even when they are not direct outputs of a hospital, they affect the way costs are transformed to outputs and should be taken into account. To uncover whether the variables have significant effect upon efficiency, we perform a non-parametric significance test. The direction of influence is retrieved from partial regression plots (see Figures A1 and A2 in the Appendix, plots for other model specifications are available upon request from the authors).

We estimate several specifications: (1) In the top panel, all specifications of Table 1.3 include all outputs (acute patients weighted by the DRG index, nursing patients, nurse/bed ratio and publications), while the bottom panel serves as a robustness check when a publication output is dropped. Effects of variables are robust across the two panels.

[17] The optimal value of m was set when the percentage of points lying above the frontier stabilized.

Table 1.3. Effects of environmental variables: whole sample

Publications output	(1) P-value		+/−	(2) P-value		+/−	(3) P-value		+/−
Nonprofit	0.052	*	−	0.012	**	−	0.040	**	−
Specialization	0.066	*	−	0.016	**	−	0.028	**	−
University	0.034	**	−	0.108	†	−	0.018	**	*NA*
2009_2010	0.110	†	−	0.048	**	−	0.062	*	−
Occupancy	0.046	**	+	0.034	**	+	0.284		+
2009_2010 × Nonprofit				<2e-16	***	−	<2e-16	***	−
2009_2010 × University				0.262		−			
No publications output	**(1) P-value**		**+/−**	**(2) P-value**		**+/−**	**(3) P-value**		**+/−**
Nonprofit	0.026	**	−	0.010	***	−	0.014	**	−
Specialization	0.096	*	−	0.056	*	−	0.056	*	−
University	0.078	*	*NA*	0.090	*	*NA*	0.064	*	−
2009_2010	0.082	*	−	0.126	†	−	0.094	*	−
Occupancy	0.024	**	+	0.016	*	+	0.002	***	+
2009_2010 × Nonprofit				<2e-16	***	−	0.004	***	−
2009_2010 × University				0.778		*NA*			

Notes: N=389; signif. codes – 0.01 '***' 0.05 '**' 0.1 '*', one-tail '†'; effects of the respective variables evaluated when all other exogenous variables are kept at the median; *NA* denotes an effect that is hardly recognizable at the median, +/− denotes the favorable/detrimental effect of an environmental variable upon efficiency; bandwidths used to smooth the kernel function are available upon request.

We found that public nonprofit hospitals tend to be less efficient than the for-profit ones, consistent with Dormont & Milcent (2012) or Czypionka *et al.* (2014). However others (Choi *et al.*, 2017; Zuckerman *et al.*, 1994; Rosko & Chilingerian, 1999; Rosko, 2001; Frohloff, 2007; Daidone & D'Amico, 2009) came to the opposite conclusion. International comparison of the effect of ownership structure on efficiency has to consider differences in the financing structure and institutional characteristics.

Regardless of ownership and legal form, all Czech hospitals are financed primarily through reimbursements from health insurance funds. Besides this, government subsidies may be provided to both nonprofit and corporatized hospitals based on the regional authority's obligation to guarantee accessibility of care in the area. Thus, the lower efficiency of nonprofit hospitals is explained by their different management structure. The result is consistent with Tiemann & Schreyögg (2012) who found that that the corporatization of German hospitals increased efficiency, even though temporarily, whereas privatization was associated with permanent increase in efficiency. The results show that the corporatization which started in 2003 was the right way to increase efficiency of regional nonprofit hospitals. Whether the effect is permanent or only temporary is subject to further research.

We observe hospitals with specialized centers to be less efficient than other units, consistent with Daidone & D'Amico (2009). Hospitals with specialized centers treat more complicated cases (average DRG index is 1.41 compared to 0.87 in non-specialized hospitals) and are more involved in research activities (average publication output 1.59 compared to 0.02). Having controlled for publications and case-mix among outputs, the fact that specialization dummy is significant suggests that the DRG case-mix index does not reflect the severity of cases properly. Indeed, the Czech DRG system was introduced as a payment mechanism in 2007 and was abandoned shortly after due to a number of drawbacks. Currently, there is a new initiative called "DRG Restart", the goal of which is to implement a new functioning DRG system before 2022. Even though not optimal, however, the DRG case mix index still decreases variation across efficiency scores of Czech hospitals when results with and without case-mix adjustment are compared.

Despite the fact that university hospitals are more involved in research reporting more publications (publication output 3.03 on average compared to 0.06 for non-university hospitals), which represents their comparative advantage relative to other hospitals in the sample, they are found to be less efficient than other hospitals. A different structure of services (more costly treatments), and management and organization of resources drive their efficiency down. Hence, even the introduction of publication output is not sufficient to make university hospitals comparable to other hospitals. The result is consistent with Rosko (2001), Grosskopf *et al.* (2001a) or Choi *et al.* (2017).

Our results thus suggest that there may be other factors specific to university hospitals and specialized centers which drive their efficiency down (e.g. they run costly research experiments, doctors' salaries may be extremely high, number of doctors may be relatively large, quality of treatment is not properly measured).

The joint dummy variable for years 2009 and 2010 reflects the introduction of user charges for each inpatient day which increased hospitals' revenues on one hand, and potential fiscal stress due to the financial crisis on the other hand. The results indicate that hospitals were not under fiscal stress that would force them to undertake restrictive measures, as the efficiency of hospitals is lower in years 2009 and 2010. Additional revenues from user charges seem to influence costs, but do not translate to outputs of our analysis. We, however, cannot say that hospitals waste more money, as these financial resources may contribute to outputs not measured in our analysis.

Occupancy rate reflects the utilization of potential capacity in a hospital. In the short term, the number of beds is given and a hospital has fixed costs related to its capacity; if the number of patients is far below a hospital's capacity, the hospital is expected to be less efficient as shown for an occupancy rate below 0.7. For rates above 0.7, the effect is not clear anymore (see the partial regression plot in Figure A1 in the Appendix) suggesting that hospitals may target below its potential capacity to accommodate fluctuations in emergency admissions (Jacobs & Dawson, 2003).[18]

In models (2) and model (3), we additionally test for the specific behavior of university and nonprofit hospitals in years 2009 and 2010. On the one hand, we do not find any significant effect for university hospitals. On the other hand, the effect for nonprofit hospitals is significantly negative. Hence, the efficiency of nonprofit hospitals decreases in years 2009 and 2010 even more relative to other hospitals and the misuse of additional financial resources in nonprofit hospitals is even more alarming.

Table 1.4 provides summary statistics of efficiency scores for the conditional efficiency model ($\theta(x, y|z)$) and unconditional model ($\theta(x, y)$).[19] The mean of both unconditional and conditional efficiencies of the whole sample is considerably high, reaching 0.933 and 0.939, respectively. Hence, a hospital can save on average 6.7% of its costs. Constrained by the operating environment, an average hospital can save around 6.1%. However, a decrease of costs cannot take place at the expense of increased re-admission rates which are then extremely costly.

Conditional efficiency analysis controls for other aspects affecting the production process and increases/decreases efficiency of a unit operating in detrimental/favorable environment, thus the standard deviation of efficiency scores for the conditional model naturally decreases.

Table 1.4. Summary of efficiency scores

	Whole sample		Small and medium		Big	
	$\theta(x, y)$	$\theta(x, y\|z)$	$\theta(x, y)$	$\theta(x, y\|z)$	$\theta(x, y)$	$\theta(x, y\|z)$
Mean	0.933	0.939	0.922	0.925	0.955	0.970
Median	1.000	1.000	0.987	1.000	1.000	1.000
Min	0.408	0.391	0.408	0.391	0.544	0.673
Max	1.323	1.043	1.323	1.043	1.101	1.006
St.dev.	0.151	0.110	0.168	0.123	0.101	0.063
Efficiency ≥ 1	211	210	130	126	81	84
Efficiency ≥ 1.1	24	0	23	0	1	0
No. obs	389	389	266	266	123	123

Note: One benchmark for the whole sample and also for the size groups. Efficiency scores may be > 1 as given by the definition of the order-m Free Disposable Hull.

[18] The essay presents only the final models which have the best fit. We also tested the effect of the number of beds, however, it was not significant. Moreover, the number of beds is highly correlated with the number of patients adjusted for DRG (correlation 0.94). We then tested the effect of cost conditions (average gross salary) on hospitals' efficiency, but no significant result was found either. All the tested specifications are available upon request.

[19] Efficiency scores for alternative specifications of the conditional model (1), (2) and (3) are very similar (Spearman's correlation coefficients vary from 0.97 to 0.988 (and are around 0.864) within (across) models with and without publications), hence we present summary of scores only from the preferred model (3) with the publication output.

We disaggregate the sample of hospitals according to size to uncover different patterns of efficiency for small and medium, and big hospitals; big hospitals treat more than 20,000 patients a year on average. Unconditional and conditional mean efficiencies are lower for small and medium hospitals, however the efficiency of small and medium hospitals varies a lot. On one hand, there are several small and medium hospitals with very low scores far below the least efficient big hospital. On the other hand, there are some medium and small hospitals with higher efficiency than the most efficient big hospital. When moving from unconditional to conditional analysis, we can observe much larger improvements in scores for big hospitals than for other units. It is because big hospitals are often nonprofit university hospitals with specialized centers and all these factors were found to be detrimental to efficiency.

To uncover whether the effects of environmental variables are specific to the size of a hospital and to provide a robustness check of the results, we carry out separate conditional analyses for two more homogeneous groups: (i) big hospitals and (ii) small and medium hospitals.

Concerning small and medium hospitals, effects in Table 1.5 are consistent with the aggregate results. The only difference is the insignificance of the joint year dummy for the specification with the full list of outputs, but it becomes significant when publication output is dropped. Hence, when small and medium hospitals are considered, additional revenues from user charges in 2009 and 2010 could have been spent on research activity, as the production of publication for some of these hospitals increased (average publication output increased from 0.033 to 0.056 between periods 2006–2008 and 2009–2010). Nevertheless, nonprofit small and medium hospitals seem to increase their spending without increasing outputs measured in this analysis as they become more inefficient in 2009 and 2010.

Table 1.5. Effects of environmental variables: big, small and medium hospitals

	Full list of outputs				No publication output			
Small and medium	(1)		(2)		(1)		(2)	
hospitals, $N = 266$	P-value	+/−	P-value	+/−	P-value	+/−	P-value	+/−
Nonprofit	0.034 **	−	0.016**	−	0.052 *	−	<2e-16***	−
Specialization	0.072 *	−	0.054 *	−	0.022**	−	0.038 **	−
Occupancy rate	0.068 *	+	0.038**	+	0.030**	+	0.070 *	+
2009_2010	0.358	*NA*	0.248	−	0.112 †	−	0.062 *	−
2009_2010 × Nonprofit			0.108 †	−			<2e-16***	−
Big hospitals	(1)		(2)		(1)		(2)	
$N = 123$	P-value	+/−	P-value	+/−	P-value	+/−	P-value	+/−
Nonprofit	0.052 *	+	0.106 †	+	0.344	+	0.320	+
Specialization	0.360	−	0.470	+	0.082 *	−	0.090 *	−
2009_2010	0.862	−	0.632	−	0.730	*NA*	0.572	−
University	0.060 *	+	0.148 †	+	0.132 †	+	0.196 †	+
Occupancy rate	0.006***	mixed	0.028**	mixed	0.196 †	mixed	0.356	mixed
2009_2010 × Nonprofit			0.878	−			0.864	+
2009_2010 × University			0.180 †	+			0.796	+

Notes: Signif. codes – 0.01 '***' 0.05 '**' 0.1 '*', one-tail '†'; effects of the respective variables evaluated when all other exogenous variables are kept at the median; *NA* denotes an effect that is hardly recognizable at the median, +/− denotes favorable/detrimental effect of an environmental variable upon efficiency; bandwidths used to smooth the kernel function are available upon request.

Results for big hospitals show a slightly different pattern. Contrary to the aggregate analysis, nonprofit and university hospitals tend to be more efficient within the group of big hospitals, although the effects are very weak and not significant for all specifications (under a stricter confidence level, the effect would be insignificant). It seems that these hospitals tend to have more publications (effects are more significant when publication output is included in the analysis) than other big hospitals (for example, university hospitals report a composite publication output of 3.05, while other big hospitals' average is around 0.28).

We cannot observe a significant effect of specialized centers when publications are considered among outputs. However, hospitals with specialized centers become less efficient when publications are dropped from the list of outputs, i.e. research activity is a relevant output requiring additional costs in these hospitals. Surprisingly, we do not find any significant effect of joint year dummy 2009 and 2010. Hence, it seems that two contradictory pressures (increase in revenues due to introduction of user charges and fiscal stress in the financial crisis) balance resulting in no specific effect. A weak effect upon efficiency can only be observed for university hospitals, revealing some minimal cost-saving measures when university hospitals are compared to other big hospitals.

1.4 Conclusion

This essay analyzed the cost efficiency of 81 hospitals in the Czech Republic during the period 2006–2010. We assessed how the operating costs, the only input in the analysis, translated to the following outputs: acute care patients adjusted for the DRG-case-mix index, nursing patients, the nurse/bed ratio, and publications reflecting research activity of a hospital.

We employed the non-parametric conditional order-m analysis. The conditional order-m approach overcomes drawbacks of the one-stage and two-stage approaches, namely separability conditions, parametric assumptions, assumptions of free disposability or convexity of the attainable set, all of which are quite restrictive (for more discussion see Bădin *et al.*, 2014).

Regarding environmental variables, we controlled for nonprofit status, the presence of a specialized center in a hospital, teaching status, and occupancy rate. We also tested whether efficiency increased or decreased in the years 2009 and 2010, when there was an important legislative change giving hospitals additional revenues through user charges. This period was, however, marked also by the financial crisis putting hospitals under fiscal stress. Additionally we include interaction terms that control for the effect financial pressures may have on nonprofit and teaching hospitals.

To uncover whether effects of environmental variables were specific to the size of a hospital and to provide a robustness check of the results, we carried out a separate conditional analysis for big hospitals and small and medium hospitals.

The mean of both unconditional and conditional efficiencies of the whole sample is considerably high, reaching around 0.935. Hence, a hospital can save on average around 6.5% of its costs. We observed that the differences of efficiency scores within the group of big hospitals are much lower than in the other group.

We found that corporatization, which was exogenous to individual hospital characteristics, had the desired effect. The results of the analysis suggest that nonprofit hospitals tend to be less efficient than their for-profit counterparts due to a different management structure. Most probably, the principal-agent problem is part of the explanation. The management of corporatized hospitals contracted and is answerable to the shareholders and the supervisory board who may exert more pressure over them than public bodies over the management of budgetary organizations. Given that, goals of for-profit institutions are better planned, measurable and monitored. For-profit hospitals are also subject to different accounting principles than their non-profit counterparts. The fact that budgetary organizations are not allowed to make profit and the earned amount must be put back, regardless of whether efficiently or not, contributes to lower scores of budgetary organizations in the analysis. However, Newhouse (1970) suggests that nonprofit hospitals have a different objective function which justifies their lower efficiency. For example, they may be situated in remote areas and thus their purpose is to guarantee access regardless of efficiency. Some wards in these hospitals are thus preserved even though they do not report a sufficient number of cases - in regions with alternative options these wards, and even hospitals, would be closed down.

We also uncovered that hospitals with specialized centers tend to be less efficient. There are two explanations at hand: (i) DRG case-mix index does not reflect the severity of treatments properly, and (ii) there are other specific factors which reduce efficiency of these hospitals (e.g. costly experiments, many doctors, high salaries of leading professionals in the field).

University hospitals were found to be comparatively less efficient. The complexity of cases and management structure reduces their efficiency even when controlling for publication output, which is much larger for university hospitals than in other units. However, university hospitals are more efficient relative to the sample of big hospitals.

Concerning the effect of years 2009 and 2010, we did not find that hospitals were under fiscal stress and were forced to save. On the contrary, due to the introduction of user charges, revenues and spending increased, but outputs did not increase equivalently, making hospitals less efficient in this period. Potential waste of this additional financial resources was even more alarming in nonprofit hospitals. Within the group of big hospitals, however, university hospitals were found to become more efficient in this period.

There are several lessons for policy-makers arising from this analysis: (i) Czech hospitals form a rather heterogenous group and when assessing their efficiency, big hospitals should be treated separately from other hospitals as their efficiency follows a different pattern; (ii) hospitals with specialized centers and university hospitals have specific characteristics which generally deter their efficiency and hence deserve special attention; (iii) the non-profit ownership status reduces efficiency of these hospitals. Particulary the last finding contributes to the current political discussions concerning the restructuring of nonprofit hospitals. This is not to say that all budgetary organizations should immediately be corporatized given their possibly different objective function, but it is advisable that some elements from the more efficient for-profit model be transferred to nonprofit hospitals. Ways to increase efficiency of non-profit hospitals serves as a motivation for further research.

Appendix

Figure A1. Partial regression plots: whole sample

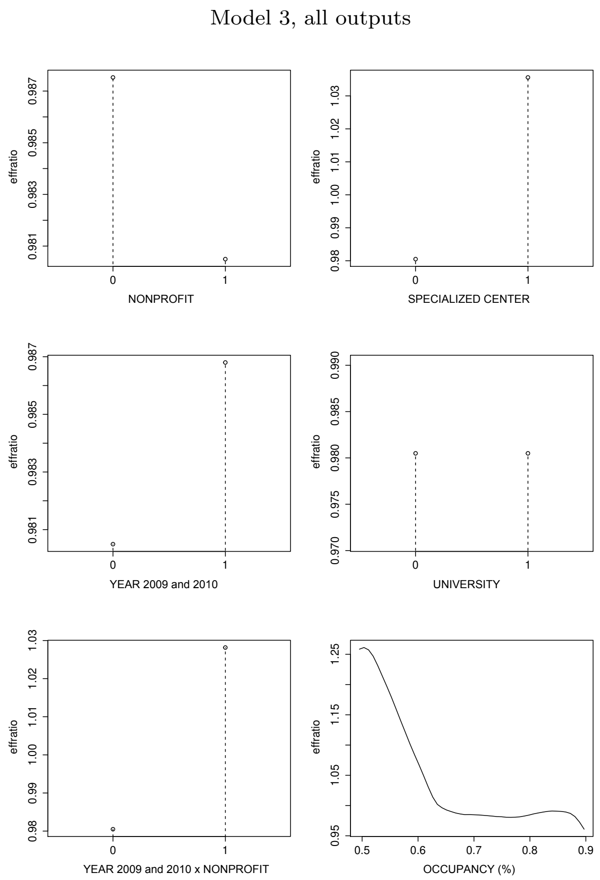
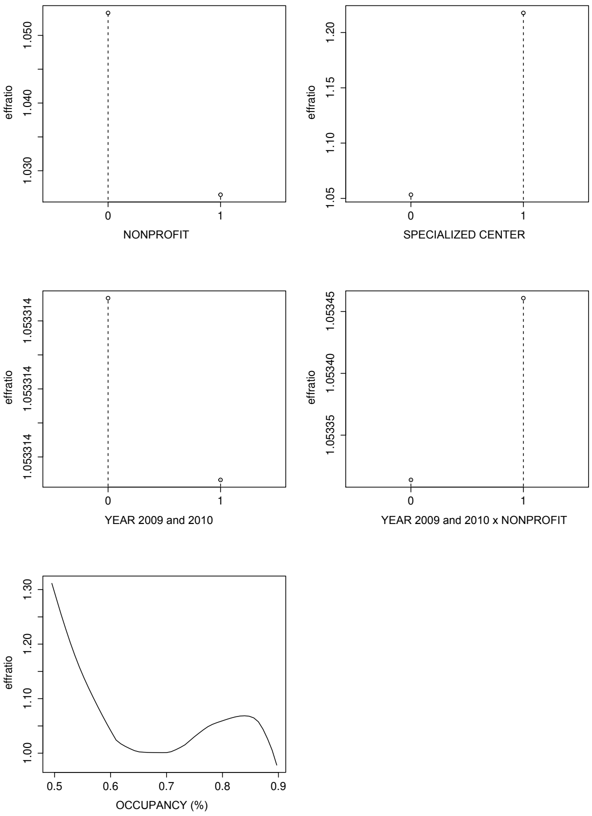
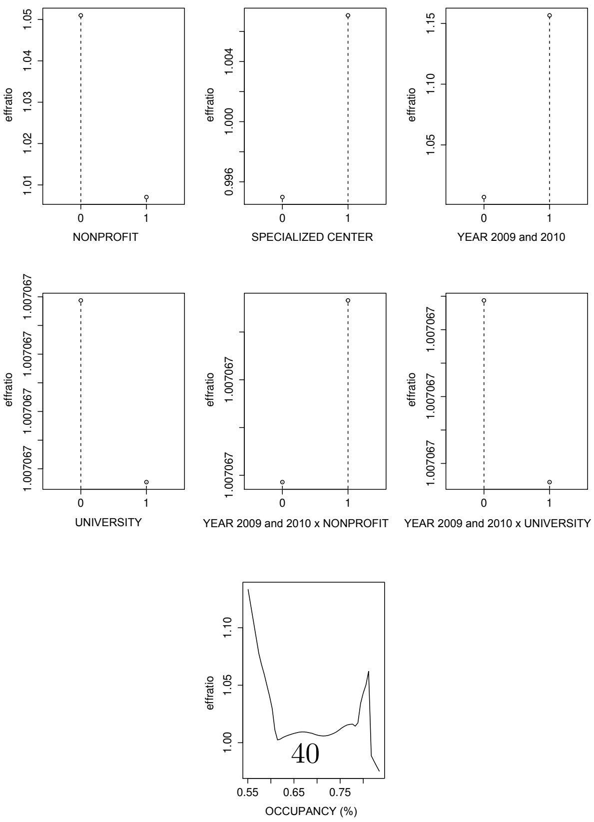


Figure A2. Partial regression plots: small and medium vs. big hospitals

Model 2, all outputs – Small and medium hospitals



Model 2, all outputs – Big hospitals



Chapter 2

The effect of inpatient user charges on inpatient care

Cost–sharing of any form in health care has been a controversial and vivid issue in all developed countries over the last two decades. Within the last 20 years cost-sharing arrangements have been changing rapidly. The definition of user charges differs across countries.[1] Within the EU, the reliance on patient cost-sharing has been increasing both in terms of scope and amount (Tambor *et al.*, 2010). A strong justification for user charges comes from fiscal pressures. Given aging population which results in increasing demand for healthcare services and escalating costs, co-payments should bring additional revenues to the system and curb the avoidable demand for healthcare services at the macroeconomic level.

However, at microeconomic level, there is a tradeoff of cost-sharing. On one hand, cost-sharing motivates the individuals to consume care efficiently. The patients forgo unnecessary care that would be consumed if they did not bear some portion of medical costs. On the other hand, some individuals - particularly the old, seriously ill and the poor - may avoid medical care that is necessary too. To avoid adverse effect on these most vulnerable groups, there is a broad equity protection in place, including ceilings or exemptions.

We may observe a correlation between patient cost–sharing arrangements and specific characteristics of the health–care and political systems (Tambor *et al.*, 2010). These include cultural values deeply rooted in the societies which cause some nations to view free healthcare as their utmost right which results in the lack of public acceptance of user charges which either restrains policy-makers from introducing user charges or results in their abolition. In countries with strong public opposition to user charges in healthcare, the attempts to introduce them depends on political representation. Their introduction is often temporary and serves the purposes of a political cycle.

[1] An overview of the latest developments and a cross-country comparison is available from the HSPM Network (https://www.hspm.org/mainpage.aspx).

Being a matter of a controversial political debate, co-payments in the Czech Republic were introduced in January 2008. Co-payments were charged for an outpatient visit during which a clinical examination was carried out (CZK 30/1.2 EUR), for a drug on a prescription (CZK 30/1.2 EUR)[2], for an inpatient day (CZK 60/2.4 EUR)[3] and for an emergency visit (CZK 90/3.6 EUR). Despite the low level of co-payments which was comparable to the price of a pack of cigarettes at the time, in February 2009, regions under the control of the Social Democratic Party (CSSD) started to reimburse patients for co-payments for care in hospitals owned by the region. It was a political decision unrelated to hospital characteristics. These benefits ceased in the middle of 2010, the exact date of cessation differed by regions.

The effect of varying levels of cost-sharing was first tested in the U.S.A. in the 1970s in the RAND Health Insurance Experiment (RAND HIE). Thousands of families were randomly assigned varying levels of cost-sharing. The resulting rich dataset has been employed for empirical analyses of cost-sharing over years (Manning, 1987; Newhouse & Group, 1993; Gruber & Kaiser, 2006) as well as for estimations of price elasticity for healthcare per se (see for example Keeler & Rolph (1988)). The RAND HIE gives even now relevant lessons for policy discussions.

Results of the RAND HIE show that cost-sharing indeed curbs demand which is in line with its primary purpose. The results are constant across healthcare services (Manning, 1987; Newhouse & Group, 1993; Gruber & Kaiser, 2006). Other studies endorse this effect analyzing different datasets. In Germany, co–payments reduced the number of doctor visits by about 10% on average (Winkelmann, 2003). An experiment conducted in the U.S.A. by Cherkin *et al.* (1990) showed that co–payments of approximately 5 USD decrease physical examinations by 14 %. According to Scitovsky & McCall (1977) the introduction of a 25 % co-insurance provision lead to approximately 24 % fewer physician visits one year later.

For a person of average health and income, the RAND HIE shows that a reasonable level of cost-sharing does not exert a negative effect on health status. Due to income-related cost-sharing and copayment ceilings, the effect on the sick and the poor was only marginally different (Gruber & Kaiser, 2006). However, other studies suggest that cost-sharing does decrease demand for healthcare of the most vulnerable groups. When estimating the effect of increased cost-sharing for ambulatory care among the elderly enrolled in Medicare plans, Trivedi *et al.* (2010) found that people in low–income and low–educated areas forgone outpatient visits most. Beck & Horne (1980) point to a similar effect for the elderly and low–income individuals in Canada. In Sweden, Elofsson *et al.* (1998) indicate that costs appeared to be the main reason to forgo a doctor's visit for 22 % of the respondents in a random sample of 17–year–olds and older. The authors link

[2] In January 2012, a single co-payment for a prescription regardless of the number of items was introduced.

[3] In December 2011, the fee increased to CZK 100/4 EUR.

this fact to their poor economic conditions since they found the probability of forgoing care to be 10 times higher among those who assessed their financial situation as poor than among those who considered it fairly good. It suggests that the demand for healthcare of these groups is very price–elastic. Zweifel & Manning (2000) explain that consumer incentives while seeking healthcare are different from a demand for consumer goods, thus the amount cost-sharing in healthcare and an appropriate exemption and ceiling plan is crucial not to worsen the health status of the most vulnerable groups.

Some studies found only a temporary effect of cost-sharing, which was sometimes offset by an increase in the level of other types of treatment, thus suggesting that there is some substitution effect at play. Roemer *et al.* (1975) showed that user charges of USD 1 for the first two doctor visits in the U.S.A. initially reduced demand for physician services, but then lead to more visits over the long–term, even more than in the control group, thus no savings resulted. Although Gruber & Kaiser (2006) found a stable effect of cost-sharing over short- and long-run in the RAND HIE, Manning (1987) showed on the RAND HIE data that a reduction in the physician services can be accompanied by increased treatment intensity in the form of longer and more expensive treatment episodes.

Saltman & Figueras (1997) and Tambor *et al.* (2010) argue that the effect of cost-sharing may depend on institutional setting and thus varies by country and the type of cost-sharing. A good example of the effect of user charges in one country is hardly replicable in another country. Some studies thus did not find any effect of cost-sharing at all, such as Schreyögg & Grabka (2008), Augurzky *et al.* (2006) or Votápková & Žílová (2016a), all of whom estimated the effect of co-payments for ambulatory services.

Most studies assess the effect of drug co-payments. A nice overview can be found in Gemmill *et al.* (2008). The effect of user charges for outpatient care was estimated for example by Farbmacher (2009) and Schreyögg & Grabka (2008). Inpatient care was assessed, for instance, in Helms *et al.* (1978).

In order to assess the effect of the mechanism, many studies exploit the fact that there are usually some exceptions to cost-sharing. Early studies, such as O'Brien (1989), employed a system of equations, taking advantage of Seemingly Unrelated Regressions. One of his equations estimated chargeable prescriptions as the dependent variable; exempt prescriptions was the dependent variable in the second regression. Some of the later studies use the difference–in–differences methodology (DiD). Examples include Winkelmann (2003), or Schreyögg & Grabka (2008) for Germany, Helms *et al.* (1978) for the U.S.A. or Zhang (2007) in Hangzhou city, China.

Czech studies analyzing the effect of user charges include Zápal (2010), who estimated the effect of user charges on children's physician visits. Physician visits are proxied by the number of drug prescriptions under the assumption that there is a fixed probability of issuing a prescription during a visit. Zápal (2010) finds a desired effect of user charges only if a one-month-period prior to the abolition of user charges (reform) is used as a pre-reform period. The results suggest a timing effect, i.e. postponement of care for the period after the reform. Kalousová (2014) estimates the effect of user charges on health service consumption among the elderly taking advantage of the SHARE database. Kalousová (2014) discovered a significant decrease in the use of outpatient care but the effect on inpatient care was insignificant. Using EU-SILC survey data, Votápková & Žílová (2016a) estimated the effect of user charges on outpatient visits finding no significant effect. A natural experiment was used in which the abolition of user charges for children represented the reform. Finally, Hromádková (2016) dealt with the effect of co-payments on prescriptions. Hromádková (2016) finds that the number of prescriptions filled decreased by 29% with the introduction of user charges. However, the effect was only temporary. The total expenditure on prescription drugs dropped only in the first quarter of the post-introduction period and then returned to the same level. A significant role was played also by a subsequent reform which allowed more packages and a different composition of drugs to be filled out on a single prescription, i.e. for a single co-payment. Finally, Hromádková (2016) analyzed behavioral responses of individuals to the partial reversal of the co-payment policy, under which patients were offered reimbursement for co-payments for drug prescription but only in region-owned pharmacies. Hromádková (2016) finds a significant preference for reimbursing pharmacies identifying also the main drivers of the preference, which include monetary costs, type of physician and distance as a proxy for opportunity costs.

The essay contributes to this stream of research and will use the difference-in-differences approach (DiD) to assess the effect of co-payments charged for an inpatient day on the amount of hospital care provided, i.e. the number of inpatient days. The advantage of the DiD is that it removes biases that could result from either permanent differences between the treatment and control groups, or shared trends.

The essay exploits the fact that the decision to reimburse user charges in 2009 was purely political and unrelated to hospital characteristics. Besides, not all hospitals within the Czech Republic are controlled by regional governments, therefore, user charge reimbursement did not apply to all hospitals within the region. There is no doubt that the patients were aware of the possibility to receive reimbursement for care in some hospitals since it was such a vivid political issue. Moreover, the hospitals were instructed to offer the possibility of reimbursement to the patients based on donation contracts or another form of agreement, thus the patients did not need to bring any money to the hospital at all, with a few exceptions in which user-charges were reimbursed with a two-months delay.

Given the socio-economic situation and political decisions at the time, we can assume that there was no confounding effect on the number of patient days other than the reform of user-charge reimbursement that would cause an expected gain bias in the DiD model for the hospitals in the treatment group (Ryan *et al.*, 2015). At the same time, the nature of the reform does not allow a "spill-over" effect from treatment to comparison group (Ryan *et al.*, 2015).

In the analysis, we are interested in the effect of user-charges on the total amount of inpatient care provided, as measured through inpatient days. Given the Czech institutional setting, Czech hospitals satisfy the demand from the community rather than they would themselves decide how much care they provide. As opposed to any other analysis of co-payments carried out in the Czech Republic which used patient-level data, a hospital will be the unit of observation in our analysis. The problem will thus be analyzed from the supply side perspective.

Similar to Hromádková (2016), the essay takes advantage of the partial reimbursement of co-payments. Regional hospitals where patients could be reimbursed for user charges in 2009 represent the treatment group, hospitals without the possibility of reimbursement constitute the control group. We control for other explanatory variables, which includes characteristics of the hospital, characteristics of the region where the hospital is situated, as well as a dummy variable acknowledging that the hospital is situated in a region where there is at least one hospital where patients could get reimbursement to account for a substitution effect.

We will answer the following questions:

1. Did the abolition of user charges for an inpatient day increase consumption of inpatient hospital care?

2. What other exogenous variables play a role in determining the number of patient days?

If the dependent variable, i.e. the number of patient days, increases after the reform which introduced user-charge reimbursement, we can conclude that the introduction of co-payments had the desired effect of reducing excess demand. Our results correspond with the mixed results of other studies carried out in the Czech Republic finding a small effect of cost-sharing. Specifically, after co-payments were reimbursed in regional hospitals, the number of inpatient days in these hospitals increased between 2.7 % and 4.1 % having controlled for exogenous hospital and regional characteristics.

The essay is organized as follows: Section 2.1 theoretically explains the methodology used, Section 2.2 introduces the dataset, Section 2.3 reports empirical results and Section 2.4 discusses the results and concludes the essay.

2.1 Methodology

We will estimate whether the reimbursement of co-payments charged on an inpatient day had an effect on the number of patient days hospitals report using a difference–in–differences (DiD) approach.

In economics, the difference–in–differences approach was fist applied in the 1980s. Ashenfelter & Card (1985) is considered a pioneering work. The simplest difference–in–differences setup is explained in Figure 2.1. The outcomes are observed for two groups of observations for two periods. One group (treatment group) is subject to the treatment in the second period, but not in the first one. The other group is not exposed to any treatment during any period. The average increase in y in the control group is then subtracted from the average increase in y in the treatment group. This reduces bias that would otherwise result from intrinsic differences if treatment and control groups were compared in the second period only, or that would result from common trends if outcomes of the treatment group were compared in isolation over time.

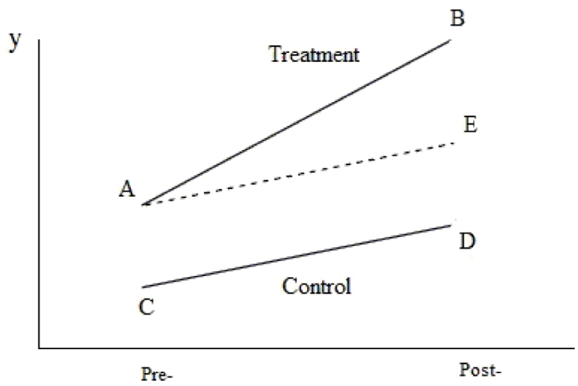


Figure 2.1. Difference–in–differences

The Difference–in–Difference model will take the following form (Blundell & Dias, 2008; Wooldridge, 2002):

$$y_{it} = \beta_0 + \beta_1 \mathbf{w}_i + \beta_2 \mathbf{z}_t + \beta_3 (\mathbf{w}_i \times \mathbf{z}_t) + \beta_4 \mathbf{X}_{it} + \epsilon_{it} \tag{2.1}$$

where y_{it} is the outcome variable, $\mathbf{w}_i$ is the treatment vector indicating whether patients of the hospital i may be reimbursed for co-payments in 2009, taking the value of 1 where reimbursement was possible; $\mathbf{z}_t$ is a vector denoting the co-payment period, thus it takes value of 1 for 2009. The interaction term $(\mathbf{w}_i \times \mathbf{z}_t)$ denotes utilization in hospitals in which patients could seek reimbursement in 2009. $\mathbf{X}_{it}$ represents a matrix of exogenous characteristics of the hospital and region that we control for.

That is, a vector of dummies, $\mathbf{w_i}$, captures the possible difference between the treatment and control groups before the possibility of reimbursement in regional hospitals. A vector of dummies $\mathbf{z}_t$ captures the aggregate factors that would cause changes in y even without a policy change. The coefficient of interest is β_3 which measures the multiplicative effect. If patients in a hospital could ask for reimbursement and the year of observation is 2009, then $(\mathbf{w}_i \times \mathbf{z}_t) = 1$.

The conditional difference–in-differences estimate, which will be obtained using OLS is then:

$$\hat{\beta}_3 = (\bar{y}_{T2} - \bar{y}_{T1}) - (\bar{y}_{C2} - \bar{y}_{C1}) \tag{2.2}$$

where C denotes control group, i.e. hospitals where patients had no choice but to pay the co-payments; and T denotes treatment group, i.e. regional hospitals where reimbursement was available.

If $\hat{\beta}_3$ is positive and significant, the reimbursement of user charges caused an increase of patient days in these hospitals. Put inversely, a positive coefficient suggests that an introduction of user charges decreases the number of inpatient days in hospitals, i.e. it reduces moral hazard and demand for inpatient hospital care. If insignificant, the introduction of user charges on inpatient day had a purely funding effect as found for example in Schreyögg & Grabka (2008) or Beck & Horne (1980).

In a significant $\hat{\beta}_3$, there may however be two effects at play. (1) user charges reduce the length of stay per patient and (2) user charges incentivize patients to move away from hospitals which effectively impose them to hospitals which reimburse user charges. That is, with a sufficiently high price elasticity, one will take advantage of care offered by regional hospitals where one receives reimbursement. In 2009, regional hospitals thus represented an outside option for patients who would otherwise consume care in hospitals that effectively imposed user charges.

However, we assume mechanism (2) to be nearly non-existent in the Czech Republic, given a low value of user charge which equals approximately the price of a lower-priced meal in a pub or a pack of cigarettes in 2009 (CZK 60). Mechanism (2) would have to be considered under a much larger user fee. Under the current set-up, the transaction costs (monetary and time) of traveling to a different hospital due to user-charge reimbursement overweight the benefits of the reimbursement itself. For instance, a train ticket from Ceske Budejovice to Plzen, which are 140 km apart, costs 140 CZK (EUR 5). The substitution effect is thus not expected to play a role given a user charge of CZK 60 (EUR 2.4), even for longer hospitalizations. In addition, when a hospitalization is longer, a patient appreciates visits and emotional support from family members whose travel costs are not imperceptible either. Travelling to a distant hospital is thus expected only for diversified treatment, which is however not a substitute to treatment in the nearest hospital.

Without the loss of generality, we assume that $\hat{\beta}_3$ estimates the national effect, user charges would have under a comprehensive reform.[4]

[4] Under the existence of mechanism (2) $\hat{\beta}_3$ would form the upper bound of the national effect.

2.2 Data

The essay analyzes a two year panel of 76 Czech general hospitals observed for the period 2008-2009.[5]

The data comes from the Institute of Health Information and Statistics of the Czech Republic,[6] Národní referenční centrum, thereafter NRC, Czech Statistical Office, thereafter CZSO, and the Registry of Companies in the Czech Republic Additional data is publicly available information of individual hospitals.

Since the Zlin region offered reimbursement for certain population groups (under 18 and above 70), and thus could not be used among the control group in the main analysis, we excluded these hospitals from the sample and used them for a robustness check only. In the robustness check of the results, we include 51 % of patient days of the Zlinsky region hospitals[7], i.e. the non-reimbursed share of patient days, as additional members of the control group. The share of 49 % of care which was potentially reimbursed was excluded fully from the analysis.

Total inpatient *days* constitute the dependent variable and are obtained by multiplying the number of patients by the average length of stay in the particular hospital.

In addition to the set of dummies introduced above to account for the treatment effect, a set of exogenous variables which are expected to affect the number of patient days of a hospital is employed. The *DRG* case–mix index accounts for the fact that people with more demanding diagnoses aim at centres which deal with more complicated cases even if outside their catchment area. The DRG case-mix index is expected to increase the number of patient days reported.

Teaching status is a dummy variable taking the value of 1 for a university hospital. University hospitals are very specific in nature. Besides treatment, their teaching and research mission is expected to increase the number of patient days they report.

Not_profit characterizes a hospital in terms of ownership status. During the process of corporatization which started in 2004, many Czech hospitals were transformed from nonprofit institutions into joint–stock companies in order to increase their efficiency. However, even the corporatized joint–stock companies are effectively under the control of districts, regions or municipalities which are their major shareholders. Only a minority of Czech hospitals is in purely private hands thus we cannot control for private vs. public ownership. The variable *not_profit* takes the value 1 when a hospital is public nonprofit and 0 otherwise, i.e. publicly controlled joint–stock company or a privately owned

[5] Year 2010 was initially considered, but it was excluded due to methodological reasons. Specifically, since reimbursement stopped in June 2010, we cannot infer much about the effect of co-payments in that year.

[6] Specifically from the following set of publications: 'Healthcare - Regions and the Czech Republic' ('Zdravotnictví kraje + ČR') for individual years

[7] The overall patient days in the Zlin region was be divided according to the shares of healthcare provided to groups 0-18 and 70+ as of UZIS (2008–2009), Table 2.13.3, *Inpatients treated in the Zlin region by age structure.*

hospital.[8]

Efficiency scores of individual hospitals obtained in Mastromarco *et al.* (2019) were applied in alternative model specifications. The variable *unc_eff2019* represents unconditional efficiency scores and *c_eff2019* are efficiency scores with efficiency conditioned on determinants. These variables are employed as robustness checks replacing the variable *DRG*. If patients perceive and value efficiency of individual hospitals, the number of patient days will increase with efficiency. Otherwise, inefficient hospitals are expected to report more patient days.

Population in the municipality where the hospital is situated characterizes the area. Besides, it is correlated with the size of the hospital since bigger hospitals are often situated in bigger cities. The population of Prague was divided into core catchment areas of individual hospitals not to bias the results. Since hospitals situated in bigger cities serve more people, the number of patient days provided is expected to increase with population.[9]

The variable *Presence* takes the value 1 if there is at least one hospital in the region where patients could ask for reimbursement of the user charge. It represents competitive pressures in the region. If price elasticity is sufficiently high, the patients will choose to travel to a hospital where reimbursement is offered. The variable would thus be negative and significant. An insignificant effect would support the initial assumption of high transaction costs outweighing benefits received from user-charge reimbursement.

Of course, the patient may travel to a hospital situated in a different region. However, in some regions the possibility of reimbursement was restricted only to the patients living in that region, thus if there was an it should be strengthened for reimbursing hospitals within regional boundaries.

All variables, including the dependent variable, except for dummies, were logarithmized due to distributional properties. Descriptive stastistics of the variables is provided in Table 2.1 and a correlation matrix is in Table B1.

2.3 Empirical results

The main results of the analysis of the effect of co-payments on inpatient care are provided in Table 2.2. Results of the models where DRG scores are replaced with efficiency scores

[8]In preliminary analysis, we initially included also a dummy for the presence of a specialized centre as defined by the Ministry of Health and a dummy reflecting whether a hospital is situated in Prague. Neither of these variables significantly increased the explanatory power of the model and both of them correlated with the variable *DRG*. Another hospital characteristics tested was the share of doctors striking in the protest "Dekujeme, odchazime" for a wage increase in spring 2010. The variable was insignificant causing a strong heteroscedasticity of the errors.

[9]The share of the elderly in a municipality was also analyzed with no significant improvement of the model.

Table 2.1. Descriptive statistics

	TREAT	POST	presence	days	DRG	teaching	population	not_profit	unc_eff2019	c_eff2019
mean	0.413	0.507	0.753	130310	1.014	0.147	71606	0.567	0.929	0.944
median	0	1	1	92503	0.875	0	24864	1	1	1
minimum	0	0	0	26785	0.650	0	3604	0	0.449	0.500
maximum	1	1	1	544025	4.220	1	371399	1	1.262	1.011
st.dev.	0.492	0.500	0.431	111000	0.400	0.354	99779	0.496	0.151	0.103

are provided in Table 2.3. Robustness checks are provided in Table B3 and Table B4.

All models were tested for normality (graphically, Jarque Bera test), the presence of homoscedasticity (Breush-Pagan test) and absence of autocorrelation of residuals (Durbin-Watson test). The test results are provided in Table B2. All models reveal autocorrelation of residuals. The main analyses with efficiency scores further report heteroscedastic errors. In addition to OLS standard errors, cluster-robust standard errors which overcome these ills are therefore reported.

The results of model 1 in Table 2.2 suggest that the interaction term is marginally significant, i.e. at 10 % significance level when heteroscedasticity and autocorrelation of residuals are addressed. The coefficient suggests that the reimbursement of co-payments increased the number of patient days in these hospitals by 2.7% after accounting for differences between treatment and control hospitals, treatment and control periods and exogenous hospital and regional characteristics. The model explains as much as 76.7 % of the variation in patient days.

Teaching and non-profit hospitals report more patient days and so do hospitals in bigger cities. The presence of another hospital in the region which offers reimbursement does not play a role when heteroscedasticity and autocorrelation are controlled for, even though the coefficient sign is as expected. Given the results of Hromádková (2016) who found that distance is a significant determinant of pharmacy choice if reimbursement is offered, we assume that the costs connected to the distance to the nearest hospital where user charges are reimbursed probably outweigh the benefits of reimbursement. This result support our assumption that the estimated effect of user charges was the true effect if user charges were applied nationally. Price elasticity at such low amount of user-charges is non-existent.

When *log(DRG)* is added into model 2 in Table 2.2, the explanatory power of the model increases and so does the effect of reimbursement. Hospitals with more complicated cases report significantly more patient days than hospitals with less complicated cases. The signs and significance of other of hospital and regional characteristics are consistent with model 1.

Table 2.2. Main results

Model 1		P-value				**Model 2**	P-value			
	Coefficient	Cluster–robust SE		OLS		Coefficient	Cluster–robust SE		OLS	
Intercept	8.8687	0.00E+00	***	1.52E-52	***	9.1464	0.00E+00	***	1.61E-51	***
INTERACTION	0.0272	8.84E-02	*	8.13E-01		0.0395	4.16E-02	*	7.28E-01	
TREAT	0.6425	1.69E-12	***	4.14E-11	***	0.6360	2.53E-13	***	3.93E-11	***
POST	-0.0224	1.93E+00		7.62E-01		-0.0212	1.82E+00		7.71E-01	
teaching	1.0690	1.83E-05	***	1.45E-18	***	0.9809	5.41E-06	***	4.79E-15	***
log(population)	0.2137	9.79E-04	***	3.55E-09	***	0.1875	2.39E-03	**	5.31E-07	***
presence	-0.2780	1.98E+00		6.04E-04	***	-0.2452	1.95E+00		2.52E-03	***
log(DRG)						0.3111	9.83E-02	*	3.31E-02	**
not_profit	0.3416	4.76E-03	***	8.33E-07	***	0.3327	6.17E-03	**	1.21E-06	***
adjusted R^2	0.7676					0.7734				
F-Statistics	< 2.2e-16					< 2.2e-16				
	142 DF					141 DF				

Note: Significance codes: 0.01 ***, 0.05 **, 0.1*

In models 3 and 4 in Table 2.3, *log(DRG)* was replaced with *log(unc_eff2019)* and *log(c_eff2019)*, respectively. The explanatory power of the models is unchanged, but none of these variables were significant. The marginal significance of the reform from models 1 and 2 in Table 2.2 disappeared. The effects of exogenous characteristics stayed consistent with the main results. The results suggest that people do not consider efficiency to be a determining factor when choosing a hospital.

Table 2.3. Results with efficiency scores

Model 3		P-value				**Model 4**	P-value			
	Coefficient	Cluster–robust SE		OLS		Coefficient	Cluster–robust SE		OLS	
Intercept	8.8292	0.00E+00	***	4.20E-51	***	8.8576	0.00E+00	***	1.52E-51	***
INTERACTION	0.0350	1.32E-01		7.64E-01		0.0452	1.30E-01		7.00E-01	
TREAT	0.6381	1.62E-10	***	1.39E-10	***	0.6283	5.85E-11	***	2.79E-10	***
POST	-0.0127	1.39E+00		8.66E-01		-0.0210	1.819		7.79E-01	
teaching	1.0563	1.57E-05	***	1.02E-17	***	1.0613	3.11E-05	***	5.93E-18	***
log(population)	0.2180	1.76E-03	***	3.13E-09	***	0.2163	2.20E-03	***	3.51E-09	***
presence	-0.2705	1.96E+00		1.04E-03	***	-0.2719	1.96E+00		9.73E-04	***
not_profit	0.3388	1.18E-02	**	1.59E-06	***	0.3314	1.48E-02	***	2.66E-06	***
log(unc_eff2019)	0.1486	5.54E-01		3.47E-01						
log(c_eff2019)						0.2041	5.98E-01		3.83E-01	
adjusted R^2	0.7669					0.7667				
F-Statistics	< 2.2e-16					< 2.2e-16				
	139 DF					139 DF				

Significance codes: 0.01 ***, 0.05 **, 0.1*

Robustness checks of the results with non-reimbursed shares of patient days of Zlin hospitals which enrich the control group are provided in Table B3 and Table B4. The effect of the reform was significant in all models except for model 1 where it turned insignificant. In other words, even though the efficiency scores stay insignificant in models 3 and 4, the effect of co-payments turned marginally significant even when efficiency scores were included (at 10 % significance level). In model 3, *log(DRG)* turned insignificant from being previously significant at 10 % level.

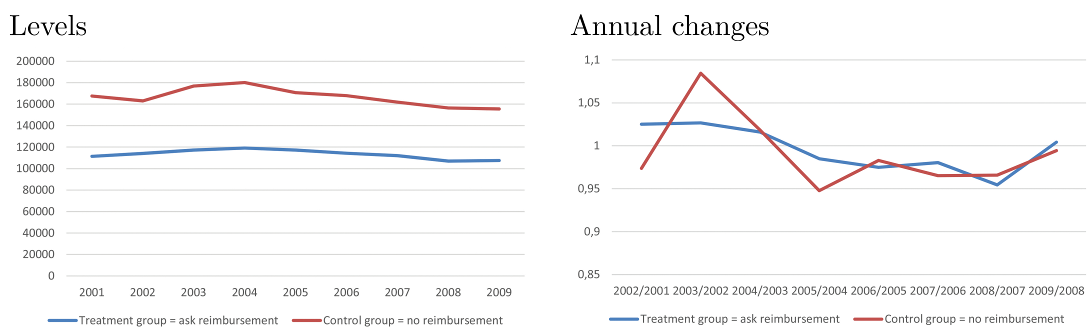
Results of an additional robustness check in Table B5 and Table B6 for which the sample was stratified by the presence of at least one reimbursing hospital in the region support the main results. However, the interaction term is insignificant in all model specifications. The effect of the reform is thus even weaker for a stratified sample than for the main analysis in which the interaction term was significant at least for some specifications. Should a substitution effect take place, the interaction term would be significant at least in the sample of hospitals with an outside option close by. This robustness check therefore points to the absence of outside options and non-existence of a substitution effect in the given set-up.

The comparison of results of the main analysis and robustness checks suggests that the effect of co-payments is very small and statistically unstable. The point estimate remains in the tight range of 2.7-4.5 % across different model specifications. Under the absence of outside options which was supported by a robustness check with a stratified sample, the observed effect is the true effect that would take place if user charges were reimbursed in all hospitals in the country.

2.3.1 Validity tests of the DiD methodology

The validity of the DiD methodology was tested. The parallel trend assumption was visually inspected. Figure 2.2 suggests that the parallel trend assumption between the treatment and control groups is satisfied. Both treatment and control groups reveal a similar trend of development prior to the examined period, i.e. 2005–2007. A slight difference occurs in the period 2003–2005 in which the number of patient days in the control group rose more relative to the treatment group. Such a development is most probably attributable to the process of corporatization of Czech hospitals and is unrelated to our purposes.

Figure 2.2. Parallel trend assumption - patient days



The parallel trend assumption was parametrically supported by a leads and lags model in Table 2.4. In order for the parallel trend assumption to hold, the effect of the lags should be insignificant in the model (Wing *et al.*, 2018). As a lead, the treatment year, i.e. 2009 was used. As lags we define dumies for years in which a parallel trend is assumed. Thus, dummies for year 2007 and 2010 were chosen as lags, even though year 2010 does not preceed the treatment. However, in 2010, the trend in the control and treatment groups should be parallel again since during the first months of 2010, regions phased out all reimbursement mechanisms, i.e. user charges were abolished for all. The fact that the entire year 2010 was not fully free from treatment may pose a minor bias that must be kept in mind when interpreting the results of the leads and lags model. The period considered for the leads and lags model was 2006–2010.[10]

The results reveal that the parallel trend assumption holds since significance of both aforementioned dummies and their cross-terms with the treatment group dummy was rejected. Even the interaction of year_2010 × TREAT was strongly insignificant despite a minor bias that may overestimate it. The explanatory variables revealed the same results both in terms of significance and the direction of the effect as the baseline model.

Table 2.4. Leads and lags model

		P-value			
	Coefficient	Cluster-robust SE		OLS	
(Intercept)	8.628	0.000	***	0.000	***
INTERACTION	0.034	0.405		0.746	
INTERACTION_2010	0.011	0.786		0.915	
INTERACTION_2007	0.033	0.392		0.756	
TREAT	0.618	0.000	***	0.000	***
POST	-0.059	1.964		0.389	
year_2010	-0.069	1.986		0.314	
year_2007	-0.001	1.028		0.990	
presence	-0.343	1.954		0.000	***
teaching	1.093	0.001	***	0.000	***
log(population)	0.265	0.003	***	0.000	***
adjusted R^2	0.724				
F-Statistics	< 2.2e-16				
	383 DF				

Significance codes: 0.01 ***, 0.05 **, 0.1*

Assuming random or exogenous assignment to treatment and control groups, the estimate of the treatment effect is more efficient with additional exogenous controls because these controls reduce the error variance. However, if the assignment is random, then including additional covariates should have a negligible effect on the estimated treatment effect. Thus, results for the treatment effect were compared for a model with additional controls and without them. The results are provided in Table 2.5. Comparing the results

[10] Dummy for year 2006 was omitted due to a dummy trap.

of the main analysis for the treatment effect, the assumption of randomness is supported since the effect on the interaction term when observables are dropped is not very much different from the main analysis when observables are included.

Table 2.5. Random assignment test

		P-values			
	Coefficient	Cluster-robust SE		OLS	
Intercept	11.5358	0.00E+00	***	4.80E-139	***
INTERACTION	0.0473	1.81E-01		8.44E-01	
TREAT	-0.0621	1.30E+00		7.17E-01	
POST	-0.0365	1.75E+00		8.14E-01	
Adjusted R^2	-0.0194				
F statistics				0.9838	
				146 DF	

Significance codes: 0.01 ***, 0.05 **, 0.1*

Since the groups differ along observables, there is a chance that they also differ along unobservables (Constatinides *et al.*, 2012), a regression of the treatment indicator on observables (a binomial logit regression) was carried out. All observables (in all models) report insignificant effect on treatment. Results are upon request from the author.

Finally a falsification test was carried out (Constatinides *et al.*, 2012). It was falsely assumed that the treatment occurred in 2006, i.e. the observed period was 2005–2006. The falsification test was carried out for model 1 (see Table 2.2) due to data availability and the fact that the treatment effect was marginally significant only in models 1 and 2 in Table 2.2. Since the variable *presence* has no meaning in the falsification test and was insignificant in Table 2.2, it was excluded for the falsification test. The results of the falsification test are provided in Table 2.6. The estimated treatment effect is statistically indistinguishable from zero, thus the observed change in 2009 likely happened due to the treatment rather than other alternative forces, even though at small levels. The size and direction of the effect of observables in the falsification test resembles the results of Model

1 in Table 2.2.

Table 2.6. Falsification test - Difference-in-differences

		P-value			
	Coefficient	Cluster-robust SE		OLS	
Intercept	8.2005	0.00E+00	***	3.42E-46	***
INTERACTION	0.0687	5.16E-01		5.77E-01	
TREAT	0.4539	6.67E-07	***	1.66E-06	***
POST	0.0112	1.10E+00		8.87E-01	
teaching	1.0102	4.74E-12	***	7.39E-16	***
log(population)	0.2706	4.60E-08	***	1.86E-11	***
not_profit	0.3312	4.84E-07	***	1.52E-05	***
Adjusted R^2	0.7652				
F-statistics				< 2.2e-16 131 DF	

Significance codes: 0.01 ***, 0.05 **, 0.1*

2.4 Conclusion

The essay estimated the effect of the reimbursement of inpatient user charges on the amount of inpatient care provided. The number of inpatient days represented the dependent variable. The analysis was carried out from the supply side perspective, assuming that hospitals respond to the demand from the community rather than themselves deciding on the amount of care they provide to the population.

The difference-in-differences methodology was applied. The decision of the social democratic regional governments of 2009 under which they decided to reimburse user charges in all hospitals under their control was unrelated to any hospital characteristics. Such a decision thus allowed a natural experiment. There was no other influence in the economy either, that would affect the number of patient days in hospitals, thus the effect of the reform was most probably plausibly exogenous bearing no expected gain bias (Ryan *et al.*, 2015). The design of the reform does not allow any 'spill-over' from the treatment to comparison groups either (Ryan *et al.*, 2015) The assumptions of the DiD, including the parallel trend assumption and random assignment, were tested in subsection 2.3.1.

As many as 76 general hospitals were observed in the period 2008-2009. The year 2008 represented the period prior to the reform, ie. control period, when all patients had to pay co-payments, and the year 2009 represented the period after the reform, i.e. treatment period in which some of the patients could be exempted from user charge payments. Hospitals under the control of the social democratic regional government which offered user-charge reimbuserment represented the treatment group.

It was assumed that if the number of patient days increases after the governmental decision to reimburse user charges, i.e. the coefficient of the interaction term denoting the treatment group in the treatment period is positive, the opposite may be inferred about the introduction of user charges. In other words, if people increase their consumption of inpatient care once user charges are abolished, they should decrease their consumption when user charges are introduced. In addition, it is believed, in given the context of very low private participation in healthcare expenses in the Czech Republic, that there is some over-consumption of healthcare present. Private participation on healthcare in the Czech Republic reached 14.3% in 2014 which is far below the European area average (WHO, 2003–2012). The number of inpatient discharges was over 21 cases per 100 inhabitants in 2016 which exceeded the Visegrad and any EU average. The number of doctor consultations exceeded 11 consultations per capita in 2013 which again significantly exceeds the EU average (OECD, 2000–2012). A reduction of the consumption due to user-charges is thus not deemed harmful, but rather as a decrease of over-consumption.

A number of specifications of the model were tested. In addition to the set of treatment dummies, we controlled for exogenous characteristics of the hospital and the region where the hospital is situated. Alternative specifications with the DRG case mix index and efficiency scores obtained in Mastromarco *et al.* (2019) were tested. A robustness check with an enhanced control group was carried out.

The alternative results are consistent as to the effect of most exogenous variables on the number of inpatient days. Teaching hospitals, non-profit hospitals and hospitals in larger cities report more inpatient days. The effect of the DRG case-mix index is marginally positive being significant at 10 % level in the main analysis but turning insignificant in the robustness check.

Having tested different model specifications, the effect of co-payments is not statistically robust. The statistical significance of the effect varies with model specification. The direction of the effect stays consistent across models. The magnitude of the effect when significant ranges from 2.7% to 4.1%.

We assumed that in the Czech set-up under a very low value of the user-charge which cost the same as a meal in a restaurant or a pack of cigarettes in 2009, the patients are unlikely to choose a hospital based on the possibility to receive user-charge reimbursement. The transaction costs to travel to a different hospital would outweigh the benefits received by user-charge reimbursement. Regional hospitals thus did not represent an outside option for patients which was also supported by the insignificant dummy variable capturing the effect of a nearby reimbursing hospital and the results of a robustness check in which hospitals were stratified by the presence of a reimbursing hospital in the region. Thus, if significant at all, user charges influenced only the length of stay per patient. The estimated effect from the natural experiment thus represents the true effect that would take place if the reform applied nationally without exceptions.

The results are consistent with existing empirical Czech literature which finds mixed evidence regarding the significance of the effect of co-payments. On the one hand, Hromádková (2016) discovered a desired effect of reimbursement in regional hospitals, but dealing with prescriptions, on the other hand Kalousová (2014) did not find any significant effect of co-payments on inpatient care using individual SILC data. Neither

Votápková & Žílová (2016a) find any significance of copayments (across age cohorts) when assessing the effect of user charges on ambulatory doctor visits. However, the level of user charge for ambulatory visits was even smaller than the user charge for an inpatient day.

Czech general hospitals analyzed in this essay treat a wide spectrum of patients. The patients are of different age, sex, income, etc. Due to universal coverage and the third-party-payer system, the structure of patients treated is, however, homogeneous across hospitals both in treatment and control groups.

Other Czech research however suggests that the effect of cost-sharing may differ for different age cohorts and type of care. Specifically, analyzing price-elasticity of different age cohorts through generic substitution, Votápková & Žílová (2016b) find out that the elderly are more price sensitive and prefer cheaper generics. When being prescribed a drug against acute illness, the patient, regardless of age, does not opt for a cheaper generic but in case of drugs against chronic illnesses, the longer the generics is available, the more the probability increases that the patient chooses the generics, although at the beginning the patients rather chose the original. It suggests that transactions costs are too high for occasional drug users, but a regular drug user recognizes benefits of saving overtime and opts for generics. It is assumed that we may observe similar results for different types of inpatient care, however for a sufficiently large amount of cost-sharing. A heterogeneity analysis of inpatient user-charges for a different structure of patients and types of care serves as motivation for further research since the present dataset does not allow for it.

Appendix

Table B1. Correlation matrix

	TREAT	POST	PRESENCE	DAYS	DRG	teaching	population	not__profit	unc__eff2019	c__eff2019
TREAT	1									
POST	-0.011	1								
PRESENCE	0.480	-0.008	1							
DAYS	-0.185	-0.009	-0.282	1						
DRG	-0.307	0.007	-0.375	0.534	1					
teaching	-0.348	-0.006	-0.287	0.856	0.530	1				
population	-0.396	0.000	-0.283	0.629	0.445	0.634	1			
not__profit	-0.359	-0.002	-0.376	0.429	0.329	0.363	0.392	1		
unc__eff2019	-0.132	-0.178	-0.129	0.113	0.134	0.112	0.003	0.017	1	
c__eff2019	-0.085	-0.115	-0.160	0.166	0.158	0.133	0.041	0.129	0.761	1

Note: Significance codes: 0.01 ***, 0.05 **, 0.1*

Table B2. Model tests

Main analysis				
Model		Normality	Homoscedasticity	No autocorrelation
	graphically	Jacque Bera test	Breush-Pagan test	Durbin–Watson test
model 1	accept	accept ** P= 0.082	accept * P = 0.015	reject P = 2.99e-09
model 2	accept	accept * P = 0.047	accept * P = 0.016	reject P = 1.624e-09
model 3	accept	accept*** P = 0.114	reject P = 0.004	reject P = 4.447e-09
model 4	accept	accept *** P = 0.134	reject P = 0.001	reject P = 4.353e-09
Robustness check				
Model		Normality	Homoscedasticity	No autocorrelation
	graphically	Jacque Bera test	Breush-Pagan test	Durbin–Watson test
model 1	accept	accept ** P = 0.073	accept * P= 0.02478	reject P = 2.272e-10
model 2	accept	accept * P = 0.03602	accept * P = 0.023	reject P = 1.313e-10
model 3	accept	accept *** P = 0.100	accept * P = 0.020	reject P = 2.926e-10
model 4	accept	accept ***P = 0.1115	accept at * P = 0.012	reject P = 3.17e-10

Table B3. Robustness results - enhanced sample: Model 1 & Model 2

Model 1		P-value				**Model 2**	P-value			
	Coefficient	Cluster–robust SE		OLS			Cluster–robust SE		OLS	
(Intercept)	8.9043	0.00E+00	***	1.73E-53	***	9.1696	0.00E+00	***	7.45E-52	***
INTERACTION	0.0277	1.06E-01		8.08E-01		0.0389	4.10E-02	**	7.32E-01	
TREAT	0.6128	8.46E-13	***	2.49E-10	***	0.6061	2.46E-13	***	2.81E-10	***
POST	-0.0221	1.91E+00		7.72E-01		-0.0210	1.82E+00		7.81E-01	
teaching	1.0693	9.73E-06	***	4.41E-18	***	0.9908	3.15E-06	***	7.97E-15	***
log(population)	0.2084	1.12E-03	**	9.10E-09	***	0.1832	3.49E-03	**	1.34E-06	***
presence	-0.2740	1.97E+00		1.00E-03	**	-0.2445	1.95E+00		3.49E-03	**
log(DRG)						0.2836	1.31E-01		5.69E-02	*
not__profit	0.3693	1.63E-03	**	1.44E-07	***	0.3623	2.21E-03	**	1.98E-07	***
adjusted R^2	0.7487					0.7531				
F-Statistics	< 2.2e-16					< 2.2e-16				
	150 DF					149 DF				

Note: Significance codes: 0.01 ***, 0.05 **, 0.1*

Table B4. Robustness results - enhanced sample: Model 3 & Model 4

Model 3		P-value				**Model 4**	P-value			
	Coefficient	Cluster–robust SE		OLS		Coefficient	Cluster–robust SE		OLS	
(Intercept)	8.8836	0.00E+00	***	3.45E-52	***	8.9010	0.00E+00	***	1.56E-52	***
INTERACTION	0.0348	8.06E-02	*	7.64E-01		0.0416	9.70E-02	*	7.22E-01	
TREAT	0.6064	2.00E-11	***	8.73E-10	***	0.6000	9.50E-11	***	1.43E-09	***
POST	-0.0158	1.49E+00		8.39E-01		-0.0212	1.84E+00		7.84E-01	
teaching	1.0622	5.55E-06	***	2.22E-17	***	1.0648	2.83E-05	***	1.46E-17	***
log(population)	0.2108	8.68E-04	***	9.93E-09	***	0.2098	1.90E-03	***	1.02E-08	***
presence	-0.2697	1.98E+00		1.49E-03	***	-0.2701	1.96E+00		1.44E-03	***
not_profit	0.3660	2.37E-03	**	3.30E-07	***	0.3608	5.76E-03	***	4.59E-07	***
log(unc_eff2019)	0.0968	6.98E-01		5.48E-01						
log(c_eff2019)						0.1446	6.91E-01		5.43E-01	
adjusted R^2	0.747					0.747				
F-Statistics	< 2.2e-16					< 2.2e-16				
	147 df					147 DF				

Note: Significance codes: 0.01 ***, 0.05 **, 0.1*

Table B5. Robustness results - stratified by presence of reimbursement in region: Model 1 & Model 2

Model 1		P-value				model 2	P-value			
	Coefficient	Cluster-robust SE		OLS		Coefficient	Cluster-robust SE		OLS	
Intercept	8.8763	0.00E+00	***	2.79E-44	***	9.3110	0.00E+00	***	4.04E-41	***
INTERACTION	0.0337	1.26E-01		7.94E-01		0.0337	1.74E-01		7.91E-01	
TREAT	0.6881	8.62E-12	***	8.44E-11	***	0.6674	3.33E-11	***	1.81E-10	***
POST	-0.0280	1.83E+00		7.70E-01		-0.0028	1.09E+00		9.76E-01	
teaching	1.5082	5.17E-11	***	5.65E-18	***	1.1876	4.12E-03	***	1.35E-07	***
log(DRG)						0.5926	2.21E-01		4.03E-02	**
log(population)	0.1836	3.33E-03	***	2.38E-06	***	0.1501	2.15E-02	**	2.58E-04	***
not_profit	0.2504	4.37E-02	**	4.98E-04	***	0.2374	5.59E-02	*	8.21E-04	***
Adjusted R^2	0.7494					0.757				
F Statistics	< 2.2e-16					< 2.2e-16				
	106 DF					105 DF				

Note: Significance codes: 0.01 ***, 0.05 **, 0.1*

Table B6. Robustness results - stratified by presence of reimbursement in region: Model 3 & Model 4

Model 3		P-value				Model 4	P-value			
	Coefficient	Cluster-robust SE		OLS		Coefficient	Cluster-robust SE		OLS	
INTERCEPT	8.8523	0.00E+00	***	7.89E-43	***	8.8648	0.00E+00	***	3.60E-43	***
INTERACTION	0.0348	2.80E-01		7.92E-01		0.0421	2.14E-01		7.49E-01	
TREAT	0.6863	1.05E-10	***	3.02E-10	***	0.6800	3.59E-11	***	4.02E-10	***
POST	0.0146	1.28E+00		8.83E-01		0.0180	1.46E+00		8.54E-01	
teaching	1.5068	3.73E-11	***	3.16E-17	***	1.5080	3.20E-10	***	1.91E-17	***
log(population)	0.1862	4.93E-03	***	2.76E-06	***	0.1859	7.27E-03	***	2.46E-06	***
not_profit	0.2407	5.25E-02	*	1.06E-03	***	0.2351	6.72E-02	*	1.40E-03	***
log(unc_eff2019)	0.0715	8.13E-01		6.66E-01						
log(c_eff2019)						0.1509	0.7828		0.5431	
Adjusted R^2	0.7473					0.7478				
F Statistics	< 2.2e-16					< 2.2e-16				
	103 DF					103 DF				

Note: Significance codes: 0.01 ***, 0.05 **, 0.1*

Chapter 3

Health status as a determinant of pre–retirement savings

Consistent with life-cycle optimization behavior, individual savings should reach their maxima at the point of retirement (Curtis *et al.*, 2015). Earlier in life individuals usually support dependent children and thus do not save much. Later in life, savings are already being used up in retirement to cover individual´s expenses. The pre-retirement cohort however increasingly suffers from health issues too due to which their health expenses escalate (Mana & Kalnicka, 2016).

The relationship between health and wealth accumulation/savings among the elderly has been widely analyzed in the literature. Smith (1999) analyzed a panel of individuals and proxied health by self-reported general health status finding that in middle and older ages new health events significantly affect household income and wealth. Adams *et al.* (2003) analyzed a US panel of individuals aged 70+ and proxied health by a dummy variable indicating whether a doctor has ever told the respondent, that he/she suffered from a particular illness. Health conditions were found to be associated with increased dis-saving from liquid wealth. Ricketts *et al.* (2013) carried out a 2SLS model analyzing the relationship between health outcomes and the willingness to save on a cross–section of individual data. They find that health outcomes are positively related to the individuals' willingness to save.[1]

Evidence show that assets do not decrease with age per se, however age changes the portfolio of assets that the households own (Guiso *et al.*, 2002; Börsch-Supan & Lusardi, 2003; Coile & Milligan, 2006; Poterba & Samwick, 1997), although earlier evidence from 1960s (Samuelson, 1969; Merton, 1969) contradicts it. Coile & Milligan (2006) observe that when aging, the households sell residence homes and vehicles, on the other hand

[1] The relationship between health status and savings has been analyzed also at the macroeconomic level. See for example Junji (2003) who found that an increase in life expectancy has a positive effect on various saving rates. Sheshinski (2009) showed that increased longevity increases aggregate savings, but only in countries with universal pension coverage and retirement incentives. The effect was found to disappear in countries with pay-as-you-go systems and high replacement rates.

increase savings held at bank accounts as they decrease the former. However, regarding primary residence, Venti & Wise (1987; 1989; 2000) and Feinstein & McFadden (1989) reject that households would sell their homes when getting older. Milligan (2005) observes that if the households do sell primary residence or vehicles, they do so very late in life.

Blundell *et al.* (2016a) and Blundell *et al.* (2016b) suggest that housing is a specific form of asset which is decumulated at much lower rate than other non-housing assets. The reasons for liquidating housing-wealth as the last are among others transaction costs, tax-advantages (in the U.S.A.) and the preference to live in own house.

Analyzing portfolio changes in terms of riskiness, Rosen & Wu (2004) and Pang & Warshawsky (2010) show that less healthy households own safer assets. The same applies for older households (Milligan, 2005). Specifically, Pang & Warshawsky (2010) show that health spending risk drives household portfolios to shift from risky equities to safer assets and enhances the demand for annuities due to their increasing-with-age superiority over bonds in hedging against life-contingent health spending and longevity risks. Atella *et al.* (2012) suggest however that health risks affect portfolio choices only in countries without a protective full-coverage of the National Health Systems (NHS) suggesting that NHS act as a shelter preventing shifts to safe investments due to individual's health issues arising in older ages.

Blundell *et al.* (2016a) find evidence and give reasons why many U.S. elderly households keep their assets into very old age and decumulate them at a lower rate than the basic life cycle model suggests. These include uncertain medical spending in an uncertain lifespan and a bequest motive, however implications for the savings patter are similar since when saving for medical expenses, the people usually die before they fully spend the savings, thus leaving bequest.

Death of a spouse as well as health shocks to either spouse were found to be very strong predictors of changing portfolio of assets after retirement, even resulting in their decumulation (Coile & Milligan, 2006), primarily regarding the sale of the primary residence, vehicles, business and real estate. Blundell *et al.* (2016a) also suggests that medical expenses, i.e. health shocks have an important effect, but primarily on non-housing assets. Poterba *et al.* (2011) add that households consider housing assets as precautionary savings drawing them down only when experiencing a shock such as a death of a partner or large medical expenses. It again supports the evidence that housing assets are sold as the last possibility. In addition, Wu (2001) finds strong differences in the behavioral pattern of the household when either partner is diagnosed with a serious illness. When a wife falls sick, general living expenses increase more than when a husband becomes seriously ill.

As the literature suggests, it is rather unlikely that specific health outcomes are exogenous to savings. The simultaneous relationship has be taken into account for instance in Adams *et al.* (2003), Meer *et al.* (2003), Salm (2010), Ricketts *et al.* (2013) or Michaud & van Soest (2004). Meer *et al.* (2003) applied an instrumental variable regression instead.

Even though the literature on the relationship between health and savings is quite extensive, it is missing for the Czech Republic. We will contribute to this stream of missing research by an analysis of the savings behavior of the Czech population aged 50–60, taking into account their health status, as well as other demographic, health and economic characteristics to account for heterogeneity in a considerably homogeneous sample. In the Czech Republic, this age cohort usually does not retire before the age of 60 although early retirement is also possible. A cross-section of Czech data of the SHARE (Survey of Health, Ageing and Retirement in Europe) database, Wave 5, is analyzed. Household savings data is normalized to individual level (divided by 2 for cohabitating partners).

The essay aims to investigate whether health status has an effect on the accumulation of quite liquid financial assets (current amount in bank accounts, the current amount saved in bonds, stocks and mutual funds and savings for long-term investment) at a certain point of time of the pre–retirement segment of the population who are assumed to be significantly concerned with savings for the years to come. At the same time, we assume that everyone wants to save some portion of their work–life income for retirement to optimize their life–time consumption behavior, consistent with the macroeconomic life–time optimization models. However, since health status worsens when older and we assume that the people tend to avert health status deterioration, among other also through investment into their own health, the effect of health status on savings should be the strongest in the pre-retirement cohort. A 2SLS methodology is applied.

In the essay, we also test for the endogeneity problem between health and the level of pre–retirement savings among the Czech population. On the one end, we assume that people with a high level of quite liquid financial assets are risk averse. Risk aversion characterizes their behavior and they are thus also sensitive to health risks. They invest into preventive health programs, go to the gym, eat healthy food, etc. which in the end increases their health status. On the other end, sicker people spend a lot of money on their health and thus save less because their primary goal is to get well now. If healthy, these consumers invest into other forms of consumption and wealth rather than health. The latter relationship has been empirically endorsed by Fu *et al.* (1999) who found that individuals who are not chronically ill are less willing to pay for low pesticide products in Taiwan compared to individuals who are chronically ill.

We will answer the following questions:

- Does health play a role in explaining the level of savings in the pre-retirement segment of the population?
- Is there an endogenous relationship between health and savings at a micro-level?
- What other characteristics play a role in explaining the level of financial assets accumulated in the pre-retirement age cohort?

The essay is organized as follows: Section 3.1 theoretically explains the methodology used, Section 3.2 introduces the dataset, Section 3.3 reports empirical results and Section 3.4 discusses and concludes.

3.1 Methodology

To account for bi–directional relationship between health outcomes and savings, we use the Two-Stage model (Greene, 2002). This approach ensures that error terms in the dependent variables are uncorrelated with the independent variables due to instrumenting.[2] A similar approach was taken by Meer *et al.* (2003) or Ricketts *et al.* (2013).

A two–stage model rests on the idea that first, two reduced form equations are estimated, such that each endogenous variable is expressed as a function of only exogenous variables (instruments) regardless of whether they have a direct effect on the respective structural equation itself or not. That is, by instrumenting both endogenous variables on all exogenous variables from both equations, we extract the endogenous part of the $y's$. Since fitted y_{hat} is then the part of y which is associated solely with the exogenous movement, i.e. changes in x and z. Excluding some elements of x from the reduced form equation would lead to loss of efficiency and consistency (or both) of the 2SLS estimator.

The obtained fitted values of the endogenous variables (health outcomes, savings) are and consequently plugged into the structural equations. For a two-stage least squares estimation (2SLS), consider the following model:

$$y_1 = \alpha_1 y_2 + \sum_{i=1}^{k} \beta_i x_i + u \tag{3.1}$$

where y_2 is an endogenous variable of a structural equation for y_1, $x_1 \dots x_k$ are exogenous independent variables and α_1 and β_1, β_k are coefficients to be estimated. In addition, there is a vector of instruments $\mathbf{z} = (1, x_1, \dots x_k, z_1 \dots z_m)$ that are correlated with endogenous y_2. In other words, vector $\mathbf{z}$ covers both exogenous independent variables and instruments. For illustration, the reduced form equation for y_2 takes:

$$y_2 = \delta_0 + \sum_{i=1}^{k} \delta_i x_i + \sum_{i=1}^{m} \gamma_i z_i + \epsilon \tag{3.2}$$

thus we get

$$y_2 = \hat{y_2} + \epsilon, \tag{3.3}$$

where $\hat{y_2}$ is a linear projection of y_2 with all exogenous variables that are not correlated with u in Equation 3.1. The projection $\hat{y_2}$ is thus not correlated with u, and ϵ takes up all the correlation with u. In a two-step estimation, $\hat{y_2}$ is then plugged into equation 3.1.

[2]A perfect instrument satisfies the lack of correlation and the exclusion condition, the latter of which however cannot be tested. Finding a good instrument is thus always a challenge.

However, in such a two-stage procedure where predicted variables are used, standard errors are usually smaller than if estimated in a single step. We thus consider only vectors **x** and **z** and estimate all in one step.

In order for the structural equations to be identified, exclusion restriction has to be satisfied (Greene, 2002), i.e. at least one variable has to be excluded from each structural equation to be identified.

3.2 Data

The dataset used for the analysis was retrieved from the SHARE database, Wave 5. We use only the Czech part of the survey of individuals aged 50-60, i.e. those born after 1951. Only individuals living with a partner are covered, i.e. single households are excluded due to homogeneity of the sample. The final dataset used for the main analysis covers 1,293 observations. A robustness check restricts the dataset to only household heads, which sum up to 742 observations.

We analyze a cross-sectional sample of individuals. Data availability and a need of a homogenous sample of the pre-retirment segment of population, disables a panel data analysis. The individuals interviewed across waves are not identical because in previous waves, the individuals born after 1951 were too young to be interviewed in previous waves.[3]

Variables used in the model shall be divided into endogenous variables and exogenous variables, the latter of which includes demographic factors, health factors and economic factors. If available, imputed variables as provided by SHARE were preferred (Börsch-Supan, 2016) due to a large number of missing observations in the original dataset.[4]

3.2.1 Endogenous variables

Household gross financial assets (‘S_G’) represents the dependent variable of the model. Is is obtained as a sum of the current amount in bank accounts, the current amount saved in bonds, stocks and mutual funds and savings for long-term investment, all for both co–habitating partners together. The resulting sum is normalized to an individual level since co–habitating partners share the costs of living and make investment decisions jointly. **Household net financial assets** (‘S_N’) normalized for an individual level was tested as an alternative dependent variable. It was obtained as household gross financial assets net of liabilities.

[3] After a modification of the research question, a panel data analysis is a motivation for further research.

[4] Except for gender, all variables used in the analysis were imputed. The variables with a low prevalence of missing values were imputed by univariate methods, such as hot-deck and regression imputations. Monetary variables with a high number of missing values were imputed using a fully conditional specification (FCS) method which preserves the original correlation structure of the variables (van Buuren *et al.*, 2006).

Self-perceived health status (‘H_{PHS}’) is an endogenous variable proxying health. It is a Likert scale[5] with values 0-5, where 0 is the worst and 5 is the best. It is assumed that if sick, one saves less, because she wants to get better as soon as possible and thus invests in her health. If healthy, individuals have a wide choice to invest in other forms of consumption and wealth/savings. It is further assumed that in the pre-retirement cohort, preference should be given to liquid financial assets, thus a positive relationship is expected similar to Ricketts *et al.* (2013). **Long–term illness** (‘H_{L-TI}’) represents an alternative proxy for health status. It is a dummy variables taking value of 1 if absent, and zero otherwise. Long–term illness may be chronic or may limit an individual in certain activities, including labor market activities, for a longer period of time. Similar as above, absence of a long-term illness and thus a better health status is expected to increase one’s pre-retirement savings.

From the dataset we cannot retrieve whether the individual had felt sick long before the interview or whether his/her health status is caused by a recent shock. Coile & Milligan (2006) show that health shocks cause changes in the financial portfolio and decumulation of wealth obviously because the sick want to pay to get well. A permanent ill-health is thus expected to disable the individuals to form savings in the first place. The inability to distinguish between a shock and a permanent status may bias the results by the fact that if the health shock happened immediately prior to the interview, the respondents may not have responded to the shock with decreased savings yet. The bias is however not expected to be large but must be kept in mind while interpreting the results of the analysis.

3.2.2 Exogenous variables

Education of the respondent (‘edu’) in years is an ordinal variable representing the level of education, i.e. human capital one has accumulated. It is assumed that more educated individuals are more concerned with their health and have a positive attitude towards savings, thus increasing both.

Income (‘income’) represents a sum of an average monthly income of a household in the year prior to the interview and an average monthly income from rents on real

[5]In the analysis the variable is treated as continuous. There has been a controversy with opponents arguing against using Likert scale as a continuous variable (for example Jamieson (2004) or Kuzon *et al.* (1996)) and proponents justifying it (for example Stevens (1946), Grolnick *et al.* (1997) or Harwell & Gatti (2001)). Estimating the first stage using probit or logit is not necessary because in 2SLS the consistency of the estimates in the second stage are not dependent upon specifying the correct functional form in the first stage (Kelejian, 1971). The more so that we are interested in the interpretation of the structural equation for savings. Likert scale may be used both as a continuous variable and a discrete choice of answer options depending on the purpose. If the researcher is interested in how the scores differ with respect to the independent variables, Likert scale is used as a continuous variable. It is generally acknowledged that likert scale be used as continuous in parametric regressions if it has at least 5 point (the more, the better as Wu & Leung (2017) show). The more points, the more the intervals between points are equal. However, a response pattern across subgroups is of interest, then likert scale is treated as a discrete variable and a log-linear or multinomial logit model would be more appropriate.

estate in the last year. It expresses heterogeneity in the earning capacity that influences how much money can be substituted between savings and other investment/consumption having satisfied basic needs. We assume that income has a positive effect on savings as found by Ricketts *et al.* (2013) or Fioroni (2010).

Assets ('assets') represents household's real assets. It was obtained as a sum of the value of main residence owned, value of own business, value of cars owned, value of other real estate, net of mortgages. Real assets represent another dimension of savings which suggests that a negative effect will result. In other words, if sold, the money would probably be saved, thus subsequently increasing liquid savings. Similarly, in the pre–retirement age, many people invest in their homes not to have to do so when old, thus decreasing liquid savings. Nakajima & Telyukova (2020) however show evidence that retired home-owners reveal a flat or increasing profile of net wealth, as opposed to other rentners. It suggests that home ownership may be a complement, rather than a substitute, to the holding of liquid assets. If the latter applies, the effect of real assets on liquid financial assets will be positive. Even though primary residence and secondary real estates and cars are expected to behave differently from the economic point of view, such that the former is a necessity, while the latter is a luxury good, aggregation is not considered to hamper the relationship anyhow.

Perceived life–expectancy ('lifex') expresses respondent's probability of living in 10 years from the interview. Note that the mean of 56 % for the whole sample reveals that on average a respondent expects to be alive 10 years later with 56 % self-perceived probability, not that that 56 % of 50-60-year-olds expected to be alive in 10 years from the date of the interview. It is assumed that a higher probability of living in 10 years increases current savings one has accumulated.

Partner's education ('edu_p') in years is included under the assumption that the level of education is a proxy for a certain type of human (economic) behavior and that partners usually influence each other's behavior.

Age ('age') is assumed to decrease health outcomes. Preliminary analyses excluded age from the savings equation, however. Since the sample covers only the pre-retirement segment, i.e. 50-60 years, it is homogeneous with respect to savings behavior. Shortly prior to retirement everyone wants to form savings. The difference of 10 year prior to retirement is too narrow to be deterministic. Preliminary analysis excluded partner's age as unimportant due to sample homogeneity. In other words, partners are usually of a similar age in this cohort (women may be a few years younger than men). Outliers characterized by enormous age differences between partners do not represent distributional issues for the analysis.

As found by Borchers & Gershwin (2012), and Whitaker *et al.* (2012) or Knolll *et al.* (2012) individual's savings plans tend to be influenced by gender. On the other hand, Ricketts *et al.* (2013) find that there is not a significant difference between males and females. Our preliminary results also suggested an insignificant effect. A variable **gender**

Table 3.1. Descriptive statistics

	mean	median	min	max	sd	n
H_{PHS}	2.87	3	1	5	1.05	1293
H_{L-TI}	0.58	1	0	1	0.49	1293
S_G	9754	5187	0	159365	12665	1293
S_N	9158	4761	-13448	159365	12844	1293
age	56.76	57	50	61	3.03	1293
gender	0.45	0	0	1	0.5	1293
edu	12.5	12	3	25	2.82	1293
income	13118	11757	0	52330	8001	1293
BMI	2	2	1	3	0.75	1293
lifesat	7.33	8	0	10	1.9	1293
assets	93313.3	61475	-13852	835390	105765	1293
lifex	55.79	50	0	100	29.34	1293
edu_p	12.57	12	3	25	2.81	1293

('gender') thus will be included only in the health equation with 1 denoting a male and 0 a female.

Body mass index ('bmi') is an indicator reducing/increasing health, rather than an indicator of poor self–control which would reduce savings (such an effect is assumed to be rather rare in the Czech context). It will thus be included only in the 'health' equation. The bmi will be divided into three categories (1) bmi$<$ 25 denoting underweight[6] and normal including 362 observations; (2) 24 $\leq$ bmi $\leq$ 30 denoting overweight including 565 individuals and (3) bmi$>$30 denoting obese with 366 observations.

Life–satisfaction ('lifesat') is used similar to Ricketts *et al.* (2013) who included job satisfaction under the assumption that psychological factors also determine one's health. This variable enters only the 'health' equation. We reject a seemingly endogenous relationship between health and life satisfaction because being healthy does not necessarily mean being satisfied with life. Healthy people may sometimes not realize the true value of life, i.e. the fact that they should be happy because they are healthy, thus they worry without a serious reason. We admit that the endogenous relationship may apply for the people who have overcome serious health issues. On average, however, we do not assume the endogenous relationship to apply. The variable takes value 0-10, with 0 being absolutely dissatisfied and 10 being absolutely satisfied and happy.

Table 3.1 reports descriptive statistics. In order to decide on the final specification of the equations, correlation conditions were checked. Correlation matrix is provided in Table C1. Correlation was insignificant among all explanatory variables.

The model was estimated in the econometric software R (R Development Core Team, 2006), packages IVREG and SYSTEMFIT and the econometric software STATA.

[6]There are only 6 observations with bmi$<$ 18.5.

3.3 Empirical results

To satisfy the exclusion restriction, from each structural equation, we exclude variables, which are only indirectly related to the particular equation, similar to Ricketts *et al.* (2013).

The structural equations take the following form:

$$S = \beta_0 + \alpha_1 \hat{H} + \beta_1 \text{edu} + \beta_2 \text{income} + \beta_3 \text{assets} + \beta_4 \text{lifex} + \beta_5 \text{edu\_p} + u \qquad (3.4)$$

$$H = \beta_0 + \alpha_1 \hat{S} + \beta_1 \text{age} + \beta_2 \text{gender} + \beta_3 \text{edu} + \beta_4 \text{income} + \beta_5 \text{bmi} + \beta_6 \text{lifesat} + u \qquad (3.5)$$

where H defines a dependent variable representing health (perceived health status, or alternatively absence of a long–term illness), S denotes gross financial assets, or alternatively net financial assets attributable to an individual. Parameter u denotes the error term.

All of the independent variables from equation 3.4 and equation 3.5 serve as instruments in the reduced form equations for health and savings.

Results of the four alternative specification of equation 3.4 are provided in Table 3.2 and Table 3.3.

Model 1 and model 3 use gross savings (S_G) as a dependent variable, whereas model 2 and model 4 employ net savings (S_N) as a dependent variable, i.e. gross savings net of household debt. Perceived health status (H_{PHS}) is an instrumented variable in models 1 and 2, while a dummy denoting absence of a long-term illness (H_{L-TI}) stands for a health status variable in models 3 and 4.

All four models reveal consistent results as to the direction and significance of exogenous variables. Sizes of the effect of the exogenous variables are not very different across models either, suggesting that selected alternative dependent and instrumented variables may be used as robust substitutes.

Additional years of education of both partners in the household increase the level of financial assets. Income also has a positive effect on financial assets accumulated as expected. Possession of real assets complements accumulation of financial assets rather than substituting it. That is, individuals who own real assets have also accumulated liquid financial assets which points to their risk-aversion. Life expectation, i.e. one's probability of living in 10 years, surprisingly decreases the level of current financial assets already accumulated, contrary to our initial expectations. Note that the cohort analyzed is very young and thus may on average still be active in the labor market without serious difficulties. It suggests that if one expects to stay alive during the 10 coming years, she probably believes that she will stay in the labor market for some more time and be able to accumulate additional financial assets for retirement.

Better health, as expressed either by self-perceived health status or by the absence of a long-term illness, increases current financial assets as initially expected. Note that the data does not allow us to distinguish between a permanent ill-health and a recent shock to health. If the health shock took place shortly prior to the interview, the individuals may not have responded to it by decreased savings by the time of the interview. Even though this bias is not expected to be large, if present, the resulting effect would underestimate the true effect. That is, without the bias, the positive effect of health on current financial assets would be even stronger.

All consistency tests necessary for the 2SLS model are reported in Table 3.4. To test **endogeneity** we applied the Wu-Hausman test which in all cases rejected the null hypothesis that OLS and 2SLS are equally consistent. Thus 2SLS here is consistent and OLS is not, therefore 2SLS is preferred. Note that if OLS is BLUE, 2SLS has larger standard errors for the coefficient estimates. Additionally, we extracted residuals from each reduced form equation. A residual from a reduced form health equation was plugged into a structural savings equation and a residual from a reduced form savings equation was plugged into a structural health equation. If the residual is significant, the variable from which the reduced form equation was extracted is endogenous. Table 3.4 reveals that health is endogenous with respect to savings which supports the results of the Wu-Hausman test.

Weak instrument test and an F-test for a joint significance of instruments which was applied on reduced form equations consistently **reject the null hypothesis that the instrument is weak**, i.e. jointly 0. In other words, results reveal that the instruments selected are sufficiently correlated with the endogenous variable.

Sargan test and the LM overidentification test corroborate **the validity of exogenous instruments.** Specifically, the Sargan test here does not reject the null hypothesis that the model is not overidentified. The LM test assumes that at least one instrument is valid and tests for the validity of all other instruments. To carry out the LM test, we regress residuals extracted from the structural savings equation on all instruments. Since the null hypothesis is not rejected, the model is not overidentified.[7]

A test for **heteroscedasticity** is essential for the 2SLS model. The Breush-Pagan test for heteroscedasticity rejects homoscedasticity at a bordering level for all four models. The White test rejects homoscedasticity for model 3 and model 4 completely. Since robust standard errors are appropriate even under homoscedasticity, Table 3.2 and Table 3.3 report also heteroscedasticity-robust standard errors. Since homoscedasticity was rejected at a bordering level only, robust standard errors do not differ much from the heteroscedasticity non-adjusted robust standard errors.

Model 2 was subject to an additional robustness check where the behavior of only

[7] An R-squared insignificantly different from 0 and an F test already reveals this, however, they assume normality. Thus, the LM test is more appropriate.

Table 3.2. Instrumental variables (2SLS) regression: Model 1 and model 2

Model 1 dependent variable: **household gross financial asssets** S_G								Model 2 dependent variable: **household net financial assets** S_N						
	Wald chi2(6) = 309.13 Prob > chi2 = 0.0000 Root MSE = 11880 No. obs = 1293				Wald chi2(6) = 259.5 Prob > chi2 = 0.0000 Root MSE = 11880 No. obs = 1293			Wald chi2(6) = 314.32 Prob > chi2 = 0.0000 Root MSE = 12199 No. obs = 1293				Wald chi2(6) = 264.51 Prob > chi2 = 0.0000 Root MSE = 12199 No. obs = 1293		
					Robust							Robust		
	coef	st. Error	p-value		st. Error	p-value		coef	st. Error	p-value		st. Error	p-value	
H_{PHS}	3479.51	1165.76	0.003	***	1373.9	0.011	**	4259.353	1193.822	0.000	***	1417.213	0.003	***
edu	353.623	154.932	0.023	**	159.41	0.027	**	360.0585	158.6616	0.023	**	162.0374	0.026	**
income	0.368	0.046	0.000	***	0.056	0.000	***	.3439394	.047553	0.000	***	.0579154	0.000	***
assets	0.009	0.003	0.011	**	0.004	0.022	**	.0099502	.0035522	0.005	***	.0039352	0.011	**
lifex	-61.902	16.021	0.000	***	17.558	0.000	***	-67.75257	16.40641	0.000	***	18.05339	0.000	***
edu_p	660.578	138.408	0.000	***	142.89	0.000	***	676.2861	141.7392	0.000	***	146.7042	0.000	***
_cons	-15166.1	2201.76	0.000	***	2235.4	0.000	***	-17741.63	2254.754	0.000	***	2324.628	0.000	***

Note: Signif. codes: 0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.'

Table 3.3. Instrumental variables (2SLS) regression: Model 3 and model 4

Model 3 dependent variable: **household gross financial asssets** S_G								Model 4 dependent variable: **household net financial assets** S_N						
		Wald chi2(6) = 272.63 Prob > chi2 = 0.0000 Root MSE = 12662 No. obs = 1293			Wald chi2(6) = 225.78 Prob > chi2 = 0.0000 Root MSE = 12662 No. obs = 1293				Wald chi2(6) = 372.99 Prob > chi2 = 0.0000 Root MSE = 13070 No. obs = 1293			Wald chi2(6) = 226.83 Prob > chi2 = 0.0000 Root MSE = 13070 No. obs = 1293		
					Robust								Robust	
	coef	st. Error	p-value		st. Error	p-value		coef	st. Error	p-value		st. Error	p-value	
H_{L-TI}	11177.83	3858.232	0.004	***	4440.006	0.012	**	12753.4	3982.575	0.001	***	4597.66	0.006	***
edu	327.86	167.8642	0.051	**	177.0305	0.064	*	348.95	173.2741	0.044	**	181.471	0.054	*
income	0.4011031	0.0471909	0	***	0.0544645	0	***	0.38513	0.048712	0	***	0.05586	0	***
assets	0.0098573	0.0036057	0.006	***	0.0040081	0.014	**	0.01134	0.003722	0.002	***	0.00411	0.006	***
lifex	-67.34865	18.12796	0	**	19.42811	0.001	***	-71.1564	18.71219	0	***	20.0226	0	***
edu_p	678.4442	145.9631	0	***	151.9684	0	***	703.13	150.6671	0	***	156.787	0	****
_cons	-11756.62	1962.326	0	***	2020.102	0	***	-13553.4	2025.568	0	***	2095.81	0	***

Note: Signif. codes: 0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.'

Table 3.4. 2SLS consistency tests

Model 1 dependent variable: **household gross financial asssets** S_G							Model 2 dependent variable: **household net financial assets** S_N					
	df1	df2	statistic	p-value			df1	df2	statistic	p-value		
Weak instrument test	4	1282	29.83	< 2e-16	***		4	1282	24.8	< 2.2e-16	***	
Wu-Hausman	1	1285	14.05	0.000	***		1	1285	22.2	0	***	
Sargan	3	NA	3.301	0.3475			3	NA	1.324	0.723		
F-test joint sign. Instr.	9	1283	44.23	< 2.2e-16	***	health	9	1283	40.53	< 2.2e-16	***	health
F-test joint sign. Instr.	9	1283	38.45	< 2.2e-16	***	savings	9	1283	40.26	< 2.2e-16	***	savings
Endogeneity test				0.738		health				0.88357		health
Endogeneity test				0.000	***	savings				0.000	***	savings
LM overident. test				0.966						0.990		
Breush-Pagan heterosc. Test	6	1286	2.035	0.058	*		6	1286	1.866	0.083	*	
White homosced. Test	2	1290	3.881	0.091	*		2	1290	1.996	0.169		

Model 3 dependent variable: **household gross financial asssets** S_G							Model 4 dependent variable: **household net financial assets** S_N					
	df1	df2	statistic	p-value			df1	df2	statistic	p-value		
Weak instrument test	4	1282	11.14	0.000	***		4	1282	10.64	1.75E-08	***	
Wu-Hausman	1	1285	12.27	0.000	***		1	1285	15.734	7.69E-05	**	
Sargan	3	NA	1.951	0.583			3	NA	1.943	0.584		
F-test joint sign. Instr.	9	1283	17.25	< 2.2e-16	***	health	9	1283	17.25	< 2.2e-16	***	health
F-test joint sign. Instr.	9	1283	38.29	< 2.2e-16	***	savings	9	1283	40.05	< 2.2e-16	***	savings
Endogeneity test				0.376		health				0.441		health
Endogeneity test				0.000	***	savings				0.000	***	savings
LM overident. test				0.990						0.990		
Breush-Pagan heterosc. Test	6	1286	2.37	0.028	**		6	1286	2.546	0.019	**	
White homosced. Test	2	1290	1.52	0.218	***		2	1290	0.5151	0.598	***	

Note: Signif. codes: 0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.'

Table 3.5. Instrumental variables (2SLS) regression: Robustness check model 2

Model 2
dependent variable: **household net financial asssets** S_N

			Wald chi2(6) = 146.54 Prob >chi2 = 0.0000 Root MSE = 12647 No. Obs = 642			Wald chi2(6) = 148.47 Prob > chi2 = 0.0000 Root MSE = 12647 No. Obs = 742	
	coef	st. Error	p-value		Robust st. Error	p-value	
H_{PHS}	3171.603	1573.896	0.044	**	2026.435	0.100	*
edu	448.8093	211.2019	0.034	**	217.9724	0.039	**
income	0.334568	0.062977	0.000	***	0.078983	0.000	***
hrass	0.012914	0.004882	0.008	***	0.005406	0.017	**
lifex	-58.3567	21.77031	0.007	***	24.12524	0.016	***
edu_p	500.1838	190.4543	0.009	***	181.0879	0.006	***
_cons	-14171.8	2951.689	0.000	***	2944.291	0.000	***

Note: Signif. codes: 0 '***' 0.01 '**' 0.05 '*' 0.1 '.'

household heads was analyzed. Such an approach reduces the sample to 742 observations, but gets rid of the potential bias which may result from normalization of financial assets accumulated to an individual level. Thus, the same amount of financial assets accumulated does not appear twice in the analysis if partners live together. Only model 2 was selected since it analyzes net financial assets (without liabilities) and uses self-perceived health status as an instrumented variable. Results of the robustness check are reported in Table 3.5. The direction of the effect of the explanatory variables support the results of the main analysis.

Consistency tests for a robustness check point that a 2SLS methodology is appropriate. Perceived health status remains endogenous to the current level of financial assets accumulated and it is appropriately instrumented within the model.

3.4 Conclusion

The essay investigated the relationship between the level of short-term liquid financial assets and a health status of the pre-retirement age cohort in the Czech Republic. A SHARE Wave 5 dataset of 1,405 individuals aged 50-60 was analyzed using a 2SLS methodology because the relationship between health and savings was endogenous. The effect of other socio-demographic and economic factors was considered. This essay is the first attempt to analyze savings behavior of the pre-retirement segment of the population in the Czech Republic.

Additional education and income were found to increase savings. Possession of real assets apparently complements the accumulation of short-term financial assets rather than substituting it. In other words, it appears that the people who own real assets always want to own a proportionate level of liquid financial assets as well. Life-expectation, i.e. one's perceived probability of living in 10 years surprisingly decreases the current level of financial assets. If one believes in an optimistic outlook of one's own life, she probably believes that she will accumulate additional financial assets, and not yet retire.

Most importantly, we find that a better health status significantly increases the level of current liquid financial assets one has accumulated in the years prior to retirement. Our micro-economic analysis should thus motivate policy-makers to give priority and carefully design preventive health programs because healthy individuals could financially secure themselves for retirement. Consistent with increasing life expectancy and the demographic prediction which will be expensive for the pension system, we suggest that a well-designed reform of the pension system which increases private participation, should ease the currently tight public pension system. In addition, if the individuals save, interest rates fall, investments rise which stimulates economic growth in aggregate. Lastly, well-designed preventive health programs save healthcare resources because prevention is often cheaper than treatment, however, prior to any intervention a carefull cost-benefit analysis is necessary (Vos *et al.*, 2010).

Appendix

Table C1. Correlation matrix

	H_{PHS}	H_{L-TI}	S_G	S_N	age	gender	edu	income	bmi	lifesat	assets	lifex	edu_p
H_{PHS}	1												
H_{L-TI}	0.548378	1											
S_G	0.084682	0.018894	1										
S_N	0.09938	0.030919	0.989704	1									
age	-0.07633	-0.08121	-0.07778	-0.06723	1								
gender	-0.0396	0.010517	0.009537	0.003435	0.115087	1							
edu	0.257691	0.164481	0.292918	0.305779	-0.05355	0.054581	1						
income	0.179052	0.067892	0.356137	0.350922	-0.05547	0.031835	0.264981	1					
bmi	-0.16077	-0.11542	-0.04828	-0.04844	0.071315	0.09988	-0.10304	-0.01745	1				
lifesat	0.340996	0.219143	0.131551	0.154073	0.067424	0.022774	0.19013	0.121251	-0.00776	1			
assets	0.178264	0.088889	0.234493	0.247874	0.006289	-0.01264	0.248646	0.275654	-0.06469	0.113595	1		
lifex	0.292299	0.220811	-0.03407	-0.02781	-0.03021	-0.04136	0.064061	0.030221	-0.11217	0.224642	0.068641	1	
edu_p	0.193037	0.110781	0.305047	0.315163	-0.01808	-0.05626	0.469558	0.238365	-0.07028	0.156511	0.240381	0.051358	1

Chapter 4

Discussion and implications for health policy

Although the quality of Czech healthcare has long been recognized as very high in international comparisons (measured by different indicators, such as life expectancy at birth, survival rates of various forms of cancer and other illnesses, mortality rates at birth and until the age of 5), the Czech Republic lacks behind other EU countries in many key efficiency and financial characteristics, including low expenses to healthcare as the share of GDP, low private participation in healthcare expenses, an excessive number of patient days, long inpatient length of stay, low number of day-care interventions, etc.[1]

In order to reach the ultimate goal of increased financial sustainability of the Czech healthcare system, OECD (2018) recommends that the Czech Republic should, among other things: (1) address the challenge of an ageing population; (2) improve hospital management; (3) reimburse hospitals based on performance rather than a flat rate; (4) strengthen the role of general practitioners in the care for chronically ill; and (5) equalize availability and quality of care across the country.

This chapter connects and discusses the results of the three preceding analytical essays that assessed specific areas of the Czech healthcare sector which either have been reformed lately or which are still worth targeting. The analytical research in this dissertation thesis, as well as our other research have a direct implication for health policy too. The implications are in line with the OECD (2018) recommendations which set increased financial sustainability as the ultimate goal for the Czech healthcare system. Each section of this chapter discusses a particular policy implication and puts our analytical research into the context of health policy. Our five implications provide empirical evidence to the OECD (2018) recommendations.

Section 4.1, stresses the need for increased inpatient care efficiency. These may increase mainly through efficient hospital management and the use DRG as a performance

[1]A thorough discussion of stylized facts and an international comparison of Czech healthcare system is provided in the Appendix.

measure and reimbursement mechanism. We compare the results of our previous research in Votápková & Šťastná (2013) on efficiency of Czech general hospitals using the Stochastic Frontier Analysis (SFA) and the results of Mastromarco *et al.* (2019), that is included in Chapter 1 of this dissertation thesis, which assesses efficiency of Czech general hospitals using a conditional FDH approach and uses an enhanced list of output variables. We shed light on inefficiency determinants and point to specific areas which need to be targeted to increase the efficiency of Czech general hospitals. We suggest that specialized and university hospitals should be financed according to a separate or adjusted scheme because they provide a different service structure, both in terms of severity and spectrum of care provided, than other hospitals.

We suggest that besides fair financing,[2] an effective DRG reimbursement mechanism which replaces a flat rate and phases out investment and operational subsidies will increase efficiency of capital investments of hospitals.

Section 4.2 suggests that in order for the Czech healthcare system to be more financially sustainable, private participation in healthcare expenses should increase. It compares the results of Votápková & Žílová (2016a) and Chapter 2 which analyze the effect of the introduction and subsequent abolition of user charges, suggesting that user charges in the Czech healthcare system were too low to show any profound effect on the consumption of healthcare. We add the results of Chapter 3 suggesting that private savings may be viewed as another form of private financing of healthcare. Since health status is a positive determinant of liquid savings, particularly in the pre-retirement cohort, the Czech Republic should strengthen the role of prevention. We argue that effective preventive programs should be in place before above-marginal levels of user charges are introduced, otherwise adverse effects of avoiding healthcare due to financial reasons, thus deteriorated health status of the population, or financial difficulties of the sick and the most vulnerable may result. At the same time, the information about the inevitability of increased private participation has to be shared with the public in a way they accept it. Simultaneously, financial literacy should increase so that people understand the necessity of precautionary savings.

4.1 Inpatient care efficiency

This section compares and reviews findings reached by two recent cost efficiency studies of Czech general hospitals, i.e. Votápková & Šťastná (2013) and Mastromarco *et al.* (2019), and applies these findings for policy purposes focusing primarily on the potential of the DRG case–mix index as a performance measure. Besides, we discuss also differences in the management structure of hospitals of different forms.

[2] For treatments of the same type and quality, hospitals would be reimbursed equally. Current system of financing is based on historical prices and a flat rate which disadvantages some hospitals.

Lessons 1 and 2: Apply DRG as a performance measure and increase efficiency of hospital management

Both Votápková & Šťastná (2013) and Mastromarco *et al.* (2019) analyzed a panel of Czech general hospitals. The period of observation and sample size differed due to given constraints - some hospitals merged into larger entities, others did not provide reliable data. Votápková & Šťastná (2013) measured cost efficiency of a panel of 99 Czech general hospitals in the period 2001-2008, while Mastromarco *et al.* (2019) analyzed a panel of 81 hospitals in the more recent period 2006–2010. As many as 78 units overlap across studies.

Both studies use frontier methods, but particular models in each of the studies differ. The variety is in line with theoretical literature that suggests that different models should be used to complement each other.

Inputs to both analyses represent total operating costs, excluding capital costs and outpatient costs of hospitals. The vector of variables used as hospital output is however different in the two studies. Votápková & Šťastná (2013) used inpatient days disaggregated into acute and nursing care and did not account for the DRG case-mix index to capture severity of acute care cases which was being developed at the time when the analysis was carried out. Mastromarco *et al.* (2019) adjusts acute care admissions for the hospital DRG–case–mix index which takes into account severity of cases treated in each hospital. In addition, Mastromarco *et al.* (2019) includes a variable representing publications among outputs to account for the fact that big and teaching hospitals devote some of their productive time to research and teaching rather than just treatment.

Both of the essays explained reasons for inefficiency or pointed at areas which should be targeted for efficiency to increase. However, the set of determinants of inefficiency is different in the two studies. The adjusted vector of output variables in the latter canceled the effect of some of the environmental variables previously used. Similarities and difference between the two studies referenced are summarized in Table 4.1, a detailed variable description is provided in Table D1 and Table D2.

When comparing inputs (overall costs) to outputs, including acute days, nursing days, doctor/bed and nurse/bed ratios, in **Votápková & Šťastná (2013)** an average hospital produced very inefficiently **(mean efficiency of 0.411)** as Table 4.2 reveals. When searching for reasons of such low efficiency scores, Votápková & Šťastná (2013) discovered that if a hospital is either larger (both in terms of the number of beds and the number of patients treated), it is nonprofit, it is situated in a municipality with a larger share of the elderly or is a university hospital, it tends to be less efficient. On the other hand, hospitals situated in bigger municipalities and in regions where also other general hospitals operate, are more efficient. **Mean efficiency, when accounting for environmental variables, reached 0.866.** Note that when determinants were included into the model, the scores became more homogeneous, i.e. standard deviation decreased.

Table 4.1. Differences between Votápková & Šťastná (2013) and Mastromarco *et al.* (2019)

Votápková & Šťastná (2013): Efficiency of hospitals in the Czech Republic Prague Economic Papers 4, 2013	**Mastromarco *et al.* (2019):** Efficiency of hospitals in the Czech Republic - conditional efficiency approach JPA 51(1), 2019
panel of 99 hospitals	panel of 81 hospitals
(78 hospitals overlap)	
2001–2008	2006–2010
parametric frontier	non–parametric frontier
inputs = overall inpatient costs	
vector of outputs	
acute days	acute patients adjusted for the DRG–case–mix index
nursing days	nursing patients
nurse–bed ratio	nurse–bed ratio
doctor–bed ratio	publish
vector of determinants	
size	specialization
nonprofit	nonprofit
university	university
share of the elderly in a municipality	year 2009 or 2010
competition	occupancy rate
population in the municipality	
	year 2009 or 2010 × university
	year 2009 or 2010 × nonprofit

The results of **Mastromarco *et al.* (2019)** discovered average **efficiency** of a sample of 81 general Czech hospitals to be **0.934 and 0.940 for the unconditional and a conditional models, respectively.** Conditioning the efficiency scores on the effect of environmental variables decreased standard deviation of scores, making them again more homogeneous across the overall sample.

The adjustment for the DRG case-mix index (Kožený *et al.*, 2010) and incorporation of publications among outputs increased the average efficiency scores remarkably, increased homogeneity of the results and **turned most of the environmental variables** used in Votápková & Šťastná (2013) **unimportant**. Neither the size of the hospital, size of the municipality where the hospital is situated, share of the elderly, or the number of general hospitals providing care in the same area played a role anymore.

Obviously, **the level of the DRG case–mix index to some extent correlates with specific demographic and hospital characteristics.** For instance, the DRG case–mix index of the group of big hospitals, which are usually situated in significantly larger municipalities with more competing general hospitals[3], strongly exceeds the DRG

[3]The share of the elderly in the municipality is comparable across hospitals in the sample.

case–mix index of hospitals of other sizes. The difference in the DRG case-mix among hospitals is caused both by the different DRG base, i.e. different hospitals tend to treat different types of cases, and differences within DRG groups, i.e. the cases hospitals treat differ also in severity, not only in the main diagnosis. For instance, heart transplant is carried out only in specialized centers/teaching hospitals, not in small hospitals. Hernia surgery is carried out both in big and small hospitals, but teaching hospitals report more complicated cases of hernia as opposed to small hospitals which treat just the simple ones. **An alternative, although very crude and imperfect, to the well-functioning DRG mechanism would thus consider local demographics, size, mission of the institution but at significantly higher costs than the DRG system.** Ferreira & Marques (2016) indeed corroborate that alternative weighting device often fulfill the same purpose but at different costs.

Unfortunately, the Czech DRG system does not reflect the severity of cases optimally yet. Indeed, the Czech DRG system was introduced as a payment mechanism in 2007 and was abandoned shortly after due to a number of drawbacks. Currently, there is a new initiative called "DRG Restart", the goal of which is to implement a new functioning DRG system until 2020.

Despite the fact that hospitals with specialized centers treat more complicated cases (average DRG index is 1.41 compared to 0.87 in non-specialized hospitals) and are more involved in research activities (average publication output 1.59 compared to 0.02) both of which is captured among outputs, the presence of a specialized center used as a determinant in Mastromarco *et al.* (2019) is detrimental to efficiency. Similarly, the effect of university status on efficiency is negative both in Votápková & Šťastná (2013) and Mastromarco *et al.* (2019) despite controlling for severity of cases and publications both of which are significantly higher for university hospitals compared to other hospitals.

University hospitals and hospitals with specialized centers may be however also characterized by a different service and management structure. Other factors are likely driving the efficiency of these hospitals down, e.g. they run costly research experiments, doctors' salaries are high, the number of doctors is relatively large, quality of treatment is not properly measured, etc. A different management structure is also responsible for the negative effect of the nonprofit ownership status.

In addition, Mastromarco *et al.* (2019) found that the more a hospital approaches its full capacity, the more efficient it is. The negative effects of the financial crisis of 2009 outweighted the positive effect of additional resources obtained from the introduction of user charges which were still partly in place in 2009 and 2010. This negative effect turned strong particularly for nonprofit hospitals. The results support the finding of our other research in Votápková & Žílová (2012), which found the effect of user charge of CZK 30 (EUR 1.2) on outpatient visits to be insignificant, and Votápková (2019) which found a small effect of CZK 60 (EUR 2.4) on inpatient care.

Hospital size revealed a strong and significant **effect on efficiency scores in**

Votápková & Šťastná (2013). It was found that the bigger the hospital is, the lower is its efficiency score both when determinants of inefficiency were accounted for or not (Table 4.2). In the baseline model - the scores ranged from 0.55 for small hospitals to 0.24 for big hospitals for the baseline model; the model with determinants reported average efficiency scores of 0.95 for small hospitals and 0.72 for big hospitals. When the resulting efficiency scores were divided into size groups, standard deviation in all groups decreased with respect to the overall sample, suggesting **homogeneity within size groups**. For the group of small and medium hospitals, standard deviation of efficiency scores decreased remarkably when determinants were included. **The group of big hospitals, including also all teaching hospitals, reported the lowest within group variation in the baseline model (0.077), but the largest variation in the model with determinants (0.121)** across groups. Note also that as opposed to the group of small and medium hospitals, standard deviation increased when determinants were included. The results, on one hand, suggest that it would be a mistake not to account for environmental variables responsible for/justifying inefficiency. On the other hand, the rise in the standard deviation among efficiency scores within the group of big hospitals in the model with determinants reveals that there are likely to be other environmental variables justifying/causing low efficiency scores which need to be searched for. Note also that the large dispersion of efficiency scores among big hospitals in the environmental model is caused by the fact that **all 11 teaching hospitals belong to this group and reach the lowest efficiency scores, not exceeding 0.7 which is the mean for the group of big hospitals with determinants.**

In **Mastromarco *et al.* (2019)**, even though not strongly, one also observes a pattern when efficiency scores are divided into size groups. However, here **efficiency scores increase with an increase in size both for conditional and unconditional models which is attributed to the incorporation of the DRG–case–mix index and publications among variables,** the lack of which previously disadvantaged bigger hospitals (Table 4.2). For the group of small hospitals, both mean and median efficiency is lower than the scores for the overall sample. The group of medium hospitals reaches the mean efficiency of the overall sample but below median scores of the overall sample. Only the group of big hospitals exceeds both mean and the median efficiency scores of the overall sample.

Standard deviation between unconditional and conditional models decreased for all size groups, consistent with the overall sample. **The most homogeneous results were found in the group of big hospitals, both for the conditional and unconditional models** contrary to the models in Votápková & Šťastná (2013) in which the group of big hospitals revealed the largest standard deviation across groups in the model with determinants and the lowest standard deviation in the baseline model due to, primarily, the presence of teaching hospitals. On the other hand, Mastromarco *et al.* (2019) observe **the largest standard deviation of scores for the group of small hospitals,** for both conditional and unconditional models, again contrary to Votápková & Šťastná (2013) which reported the lowest standard deviation for small hospitals in the model with determinants and the highest standard deviation in the baseline model.

The remarkable increase in average efficiency scores for big hospitals is explained by the improved vector of outputs in Mastromarco *et al.* (2019). It suggests that **big and teaching hospitals indeed tend to treat more complicated cases and carry out research besides treatment, which is the main function of hospitals.** Also other hospitals (small and medium ones) that do research and publish their findings differentiate from those that just treat, which is supported by the large standard deviation of efficiency scores among the group of small hospitals in Mastromarco *et al.* (2019).

The results showed that corporatized hospitals reveal different management structure than budgetary organizations with corporatized hospitals being more efficient. Most probably, the principal-agent problem is the explanation. The management of corporatized hospitals is answerable to the shareholders and the supervisory board who may exert pressure over them more efficiently than public bodies over the management of budgetary organizations. However, budgetary organizations may have a different objective function too (Newhouse, 1970) which justifies a portion of their inefficiency. For instance, they may keep wards that do not provide a profitable amount of care in remote areas. Such wards would, if a part of for-profit units, most probably be closed down. We thus suggest not to corporatize all budgetary hospitals but overtaking some efficient elements from for-profit institutions would be desirable.

A robustness check in Mastromarco *et al.* (2019), where the sample was divided into two subgroups, i.e. big hospitals; and small and medium hospitals, supports the results of the main analysis with some exceptions. Concerning small and medium hospitals the only difference is the insignificance of the joint year dummy for the specification with publication output, but it becomes marginally significant when publication output is dropped. It suggests that research activity of small and medium hospitals in 2009 and 2010 is responsible for additional costs (average publication output of small and medium hospitals increased from 0.033 to 0.056 between periods 2006-2008 and 2009-2010). Nonprofit small and medium hospitals increased their spending without increasing outputs in 2009 and 2010 rather than taking cost-saving measures.

Contrary to the aggregate analysis, **nonprofit and university hospitals tend to be more efficient within the group of big hospitals when publication output is included** suggesting that indeed university hospitals report more publications than other hospitals (3.05 for university hospitals vs. 0.28 for other big hospitals). Big hospitals with specialized centers become less efficient when publications are dropped as an output pointing that **publications is a relevant output that justifies additional costs.** Contrary to small and medium hospitals and other big hospitals, university hospitals reveal some minimal cost-saving measures in 2009 and 2010

Table 4.2. Summary of efficiency scores

	whole sample				small			
	Votapkova & Stastna (2013)		Votapkova & Stastna (2019)		Votapkova & Stastna (2013)		Votapkova & Stastna (2019)	
	baseline	with determinants	unconditional	conditional	baseline	with determinants	unconditional	conditional
min	0.112	0.501	0.522	0.638	0.373	0.982	0.522	0.638
max	0.931	0.997	1.273	1.000	0.914	0.997	1.123	1.000
mean	0.411	0.863	0.934	0.940	0.590	0.993	0.917	0.915
median	0.379	0.866	0.962	0.981	0.558	0.993	0.950	0.942
st. dev.	0.192	0.133	0.128	0.077	0.145	0.004	0.158	0.099
effficiency ≥ 1	NA	NA	27	16	NA	NA	11	4
effficiency ≥ 1.1	NA	NA	3	0	NA	NA	2	0
effficiency ≥ 1.2	NA	NA	2	0	NA	NA	0	0
hospitals	99	99	81	81	33	33	28	28
	Medium				Big			
	Votapkova & Stastna (2013)		Votapkova & Stastna (2019)		Votapkova & Stastna (2013)		Votapkova & Stastna (2019)	
	baseline	with determinants	unconditional	conditional	baseline	with determinants	unconditional	conditional
min	0.112	0.809	0.655	0.785	0.113	0.501	0.640	0.835
max	0.931	0.982	1.273	1.000	0.379	0.898	1.035	1.000
mean	0.399	0.875	0.930	0.939	0.243	0.722	0.956	0.971
median	0.379	0.863	0.943	0.961	0.260	0.789	0.983	0.983
st. dev.	0.153	0.038	0.127	0.069	0.077	0.121	0.080	0.038
effficiency ≥ 1	NA	NA	6	4	NA	NA	10	8
effficiency ≥ 1.1	NA	NA	1	0	NA	NA	0	0
effficiency ≥ 1.2	NA	NA	1	0	NA	NA	0	0
hospitals	33	33	28	28	33	33	25	25

Lesson 3: Apply DRG as a reimbursement mechanism with some specifics

Research on the DRG started in the second half of 1990s, advanced substantially with the foundation of the National Reference Center (NRC) in 2003, however, the system was not implemented as a payment mechanism until **2007**.

After a short period of partial reimbursement of inpatient care through the DRG, in 2014, the DRG became again a measure of performance only and payments were based on historical prices due to multiple drawbacks of the DRG system. Currently, if a hospital fulfills certain production criteria, e.g. 97 % of inpatient production of the reference year measured through the DRG, it receives approx. 111.2 % of the budget of the reference year. (The percentages are defined in the Reimbursement Decree issued annually by the Ministry of Health of the Czech Republic and are subject to annual changes).

In 2014, DRG responsibilities of the NRC were transferred to the Institute of Health Information and Statistics of the Czech Republic (further 'UZIS') and a new phase of implementation of the DRG system started with a new project called **"DRG Restart"**. The goal of the project was to implement a new DRG system, **CZ-DRG until 2020,**, which should **correct all shortcomings** of the abandoned IR-DRG system. Based on the latest developments within the project, however, the implementation as a reimbursement mechanism is likely to be postponed to 2022.

Only acute inpatient care will be financed through the CZ-DRG. Other hospital services, such as aftercare treatment, rehabilitation, outpatient care and day cases, will still be reimbursed through per day payments or fee for service payments.

All procedures performed, drugs, blood and medical devices used, external laboratory tests outsourced, capital, operating and overhead costs expended - are covered in the cost of the case to calculate DRG relative weights. Operating and overhead costs are allocated to cases in the amount of per diem costs. Teaching, research and development costs are financed through public funds and are excluded from the DRG.

Within the DRG system, the payment for the same type of care of the same quality should be set equal, regardless of the length of stay. However, costs rise with every additional inpatient day for which a provider would no longer be reimbursed. The main potential of the implementation of the DRG is thus to **shorten the excessive length of stay**, which should lead to an **increase in efficiency** without a decrease in the quality of healthcare, holding day-case substitution and re-admission rates constant.

Reduction in the average length of stay is one of the positive effects of the introduction of the DRG reimbursement system abroad. Cheng *et al.* (2012) in their study about impacts of DRG payments on health care in Taiwan determined that the introduction of the DRG payment system significantly decreases the length of stay. Louis *et al.* (1999) found the same effect in Italy - the mean length of stay decreased from 9.1 days to 8.8 days. Moreover, Italian hospital admissions decreased as well. Even though the decreasing average length of stay is consistent with a general trend (Rosenberg & Browne, 2001) that applies also to the Czech Republic[4], further decrease of the average length of stay due to the DRG introduction is expected.

Under the optimal DRG reimbursement setup a flat base payment rate should apply to all hospitals. However, given the specifics of the highly diversified care in university hospitals and highly specialized centers, which we showed also in Section 4.1, **"risk corridors" or an alternative set of weights should be applied in these cases to reimburse the highly diversified care.**

An effectively functioning DRG reimbursement system should also increase efficiency of capital expenditures of hospitals and reduce corruption involved in capital purchases. Corruption in capital purchases is characterized by excessive price which the hospital pays, although a profoundly cheaper unit is available in the market. Hospital management is then rewarded by the seller for buying the overpriced capital.

Until April 2014, there were no standards or regulations for hospitals on how to finance and purchase individual pieces of equipment. It was the responsibility of the health care provider to decide what to buy. Hospitals themselves then contracted with health insurance funds upon reimbursement of treatment using the equipment bought.

In April 2014, the Equipment Commission was established, which assesses medical and economic efficiency of "hard" equipment[5]. The Commission analyzes usefulness of the equipment with respect to the entire Czech Republic or region, and also availability of after–care treatment neccessary after the intervention within the particular hospital and immediate neighborhood hospitals. The Commission calculates how much the intervention using the equipment will cost the healthcare system as a whole and what benefits it will bring both to the patient and the system.[6] **The Commission then either recommends or disapproves purchase of the equipment.** Health insurance funds agreed to respect the decision of the commission[7] and not to reimburse procedures carried out with equipment disapproved by the Commission. **Establishment of the Equipment Commission thus prevents over–priced, inefficient and unnecessary purchases.**

Thus together with the the Equipment Commission, if a hospital receives a fixed payment per case within the DRG reimbursement mechanism, which includes also capital costs, **hospital management is** disincetivized from corruption and is **motivated to negotiate the best possible price.** In other words, if an overpriced machine is bought, either it is not recommended by the Equipment Commission and thus should not be reimbursed by insurance funds, or if it is reimbursed, the excess costs is absorbed by the

[4]The average acute care length of stay has been decreasing continuously over the past years - in 2009, it was 7.1 days which decreased to 6.8 in 2011 (OECD, 2000–2012).

[5]These include magnetic resonance imaging, computer tomography, linear particle accelerators, angiographic lines, etc. It does not include frequently used equipment such as EKG, etc.

[6]Although, it is not an HTA analysis.

[7]Although health insurance funds are not legally obliged to respect the decisions of the Commission.

hospital itself, rather than by the insurance company if reimbursed through the case–mix mechanism. In the future, when DRG reimbursement mechanism effectively functions and the Equipment Commission is well–rooted within the healthcare system, corruption in capital purchases should significantly diminish. That is, if the price of the equipment is included in the price of service that providers produce, i.e. DRG reimbursement, health care providers should have an **incentive to negotiate the best price** in order to save money.

At the current set-up besides reimbursement from insurance funds, **state, regional, municipal and city (both budgetary and corporatized) hospitals receive investment, and sometimes operational, subsidies**; however, private hospitals do not receive any additional money. Under such a set-up with DRG in place, public hospitals would thus be reimbursed for these costs twice. The final solution of double financing of capital costs is not yet clear, however, the plan is to **abolish investment subsidies to public hospitals so that public hospitals are not reimbursed twice**.

4.2 Private participation in healthcare expenses and the role of prevention

A low level of private participation in healthcare expenses and a large share of healthcare costs which are financed publicly threaten the sustainability of future financing of the Czech healthcare system given the challenge of population aging.

The introduction of user charges has two purposes: (1) bring additional financial resources to the public healthcare system and (2) curb excess and unnecessary demand for healthcare services. The international comparison of characteristics of the Czech healthcare system, such as the length of stay, number of hospitalizations and doctor visits, etc. and a set-up of the Czech system as being almost entirely publicly funded, point to the existence of excess demand within the Czech healthcare system.

In 2008, a user-charge reform introduced fees for outpatient visits during which a clinical examination was carried out (CZK 30/1.2 EUR), for a drug on a prescription (CZK 30/1.2 EUR), for an inpatient day (CZK 60/2.4 EUR) and for an emergency visit (CZK 90/3.6 EUR). An extensive discussion both among politicians and the public followed. As a result, regions governed by Social Democracy reimbursed patients for user charges for an inpatient day in the hospitals with their influence between February 2009 and mid-2010. In addition, user charges for outpatient visits for children were abolished in April 2009. Such a set up enabled a natural experiment in Votápková & Žílová (2016a) and Chapter 2. A comparison of the results of the respective studies implies that private participation on healthcare expenses was too low to reveal any significant effect, and should be increased. We also suggest that increased private participation in healthcare expenses has to be accompanied by a strengthened role of prevention in order for it not to have an adverse effect.

Lesson 4: Increase private participation in healthcare expenses

Votápková & Žílová (2016a) analyzed a EU-SILC (Survey of household income and living conditions) micro level data of 1,841 individuals. We applied the Zero-inflated negative binomial model. The number of physician visits in the last 12 months of an individual represented the dependent variable. Independent variables included individual characteristics determining the probability of increased utilization of healthcare services, such as sex, the number of household members and annual household income per member of a household.

Chapter 2 assessed the effect of the possibility of user charge reimbursement on the number of inpatient days of hospitals. We approach the problem from the supply side perspective with hospitals being the units of observations. We assume that hospitals respond to the fixed demand from the community rather than themselves deciding on the amount of care they provide. The total of 76 hospitals are analyzed. The possibility of user charge reimbursement in some hospitals proxied user charge abolition. We assumed that the effect of user charge abolition may without the loss of generality be inversed to the reform of user charge introduction. Two effects of the reform were considered. (1) The reimbursement of user charges increases the length of stay in the reimbursing hospitals (intensive margin); and (2) under sufficient price elasticity, user charges incetivize patients to move away from hospitals that effectively impose user charges to reimbursing hospitals (extensive margin). However, under the current set-up, only intensive margin is at play because transaction costs (monetary and time) to travel even the nearest reimbursing hospital exceed the low amount of user charge. Thus the observed effect equals the national effect without outside options.

The observed period in both studies was 2008-2009, with 2009 being the treatment period, i.e. the period after effective user charge abolition. Treatment groups represent those observations which were subject to user charge exclusion/reimbursement, i.e. children in the former and hospitals with major political influence of the Social Democracy in the latter.

Votápková & Žílová (2016a) found no significant effect of the abolition of user charges for children on the utilization of outpatient care. Specifically, no effect was found on the probability of avoiding healthcare if sick either on the frequency of doctor outpatient visits. A number of robustness checks corroborate the results.[8]

Chapter 2 found only a small effect of user charge abolition on the utilization of inpatient care. The reimbursement of co-payments increased the number of patient days within the range of 2.7% - 4.5 % depending on the model, being however significant at 10%

[8] A sample of adults was restricted to (1) 18-26 olds; (2) to childless individuals, (3) up to 65 year of age; (4) employed adults up to 65 years of age.

level only. In two alternative specifications, the effect of the reform lost significance. In the robustness checks of the results with a larger control group, significance was lost in the two previously significant models, while in other two previously insignificant specifications, the effect of the reimbursement of user charges became positively significant, however, at 10% significance level only. A robustness check on a subsample of hospitals with a nearby outside option (at least one reimbursing hospital in the region) pointed to the absence of extensive margin, i.e. the absence substitution between reimbursing and non-reimbursing hospitals under the 60 CZK/2.4 EUR user-charge.

The comparison of studies assessing the effect of user charge abolition suggests no profound effect of the reform. The results may be without the loss of generality applied to the reform of user charge introduction. The result is consistent with other research of the effect of Czech user charges which found only a temporary or an insignificant effect (Hromádková, 2016; Zápal, 2010). The insignificant effect of the reform on the demand is thus most likely caused by a low amount of copayments rather than the non-existence of unnecessary demand. In other words, the amount of copayment was set at a marginal level. **With a larger copayment, the effect may be significant.** The amount of 30 CZK/1.2 EUR for an outpatient visit and 60 CZK/2.4 EUR for a hospitalization day did not sufficiently motivate the people to carefully consider the necessity of their need to visit a physician or be hospitalized. They revealed no price elasticity at these marginal amounts. At larger amounts, price elasticity may cause a different behavior. Only Kalousová (2014) noted some price elasticity to user charges among the elderly (50+). She found a decreasing significant effect on the number of doctor visits for this cohort. We assume, however, that the elderly behave differently due to their larger price sensitivity relative to other age cohorts as shown by Votápková & Žílová (2016b) who found a strong preference for cheaper generics among the elderly compared to more expensive originals. The results for the elderly thus cannot be generalized for the whole population. The results of Kalousová (2014) also suggest that **the elderly are the most vulnerable group for whom a copayment ceiling always has to be considered.**

Lesson 5: Strengthen the role of prevention

In Chapter 3, we assess the effect of health status on individual liquid savings, concentrating on the pre-retirement cohort. We analyze a sample of the total of 1,405 Czech individuals at ages 50-60 years interviewed for SHARE. We point to the endogenous relationship of health status with respect to savings, thus we apply a 2SLS methodology. Testing for alternative model specifications, we consistently find that a better health status increases current individual liquid savings in the years prior to retirement.

We assume that prior to retirement the individuals tend to secure themselves for the years when they are not active in the labor market anymore. Thus, at this age cohort, we perceive individual liquid savings to be a private source which could potentially be used for the funding of private healthcare expenses in the years to come. A positive and significant effect of health status on individual's liquid savings reveals that if healthy, the individuals save, but if not feeling well, they invest into their health status to feel better. **The government should thus strengthen the role of prevention so that the individuals accumulate liquid savings** which they would use while retired. These implication can therefore be generalized for the public pension system too. That is, **if the individuals are healthy prior to retirement, they would accumulate wealth which eases the tight public healthcare and pension budgets afterwards.** This implication is topical particularly in the view of the challenge of aging population.

We can generalize these implication for other economically active age cohorts too, although the effect is expected to be the strongest for the pre-retirement segment who anticipate years during which further wealth accumulation will be difficult due to economic inactivity. Besides, individual savings have macroeconomic implication for investments and growth.

Considering implications of Votápková & Žílová (2016a) and Chapter 2 that the amount of user charges should be above the marginal level to have an effect in decreasing excess healthcare demand, personal liquid savings, and thus prevention, have a far reaching importance both at microeconomic and macroeconomic level. Assuming that user charges would be introduced, the individuals of all ages would have to create precautionary savings for the case of illnesses, injuries or other health disorders which they can hardly predict. Without any precautionary savings, a need for clinical examination or hospitalization may cause either financial troubles of the individual or may result in avoiding healthcare for financial reasons, both of which the opponents of user charges advocated as legitimate reasons for user charge abolition.

If the system of prevention functions well in the Czech Republic, we expect also the user charge reform to have the desired effect of reducing excess demand and increasing the private share of funding of healthcare without the adverse effects of avoiding healthcare for financial reason, thus deterioration of health status of the population or financial problems resulting from the absence of precautionary savings. Naturally, financial literacy should increase over time. **The population has to learn about the importance of precautionary savings for cost-sharing and healthcare expenses from early childhood** because particularly lower educated segments of population often do not accumulate any liquid financial reserves. In addition, the politicians should unite over the necessity of user charges and should communicate it to the public in a way that they accept it. Public discontent and a loss of political preference were, after all, a reason for the terminal abolition of all user charges, except for emergency visits, in January 2015.

4.3 Conclusions

The chapter discussed the results of the essays analyzed in this thesis and put them into the context of our overall research activities. We identified five policy implications resulting from our research. The policy implications of our research are in line with OECD (2018) recommendations, the ultimate recommendation of which is increased financial sustainability of the Czech healthcare system.

Results of our research suggest that the Czech Republic should (1) increase inpatient care efficiency and (2) increase private participation on healthcare expenses in line with the strengthened role of prevention.

We suggest that inpatient care efficiency may increase if (i) **DRG is applied as a performance and reimbursement mechanism** and (ii) efficiency of hospital management increases. It is however likely that that efficiency of hospital management increases once the DRG reimbursement mechanism is applied.

Even though the DRG system used in our analysis was not optimal yet, we showed that it significantly **improves benchmarking of hospitals.** When not accounting for the DRG case–mix index in the efficiency analysis, hospitals providing highly specialized care which consumes more resources would be disadvantaged. On the other hand, efficiency of hospitals which produce just general care but consume more resources than necessary would be inflated.

In other words, not accounting for the DRG–case–mix index and publication outputs would overstate inefficiency of big and teaching hospitals, relative to small and medium hospitals. Big and teaching hospitals often treat complicated cases and besides treatment, devote a portion of their productive time to research. As a result in Votápková & Šťastná (2013), big and teaching hospitals were least efficient.

However, in Mastromarco *et al.* (2019) when DRG adjusted and publication outputs were considered, big and teaching hospitals were most efficient relative to the rest of the sample. Comparison of the two studies suggests that **big and teaching hospitals should be funded separately from small and medium ones.** If funded based on the same scheme, there will always be one group disadvantaged.

Comparison of the results of the two analyses in Section 4.1 implies that **an alternative to the DRG system would look at local demographics, size, mission and legal form of the institution** to correct funding which can be directly captured through the case–mix approach. **DRG system thus bears lower administrative costs in the long run**, once fully and effectively functioning.

The disaggregated analysis further suggests that **small and medium hospitals may potentially benefit from corporatization.** However, a deeper investigation of the setting of each hospital would be desirable before corporatization could be set as a general goal. In addition, **big hospitals were more financially flexible** than small and medium ones in their present set-up.

Effective functioning of the DRG as a reimbursement mechanism, where capital costs are included, together with the Equipment Commission should also guarantee effective capital purchases of Czech hospitals. Such a setup would disincentivise mangers from corruption and over-priced capital purchases and would increase their efficiency.

Another stream of our research implies that **private participation on healthcare expenses should increase.** Not only is the private participation on healthcare expenses in the Czech Republic extremely low compared to other countries but in the view of population aging, pure public financing of healthcare expenses is likely to be unsustainable.

We discussed the results of Votápková & Žílová (2016a) and Chapter 2 both of which analytically assessed the reform attempts to introduce user charges in healthcare, but in different sectors of healthcare. Votápková & Žílová (2016a) assessed the effect of outpatient user charges, while Chapter 2 analyzed the effect of user charges in the inpatient sector. None of the analyses found any profound effect of a decrease of excess demand for healthcare services in the Czech Republic. We assume that the level of user charges was too low to show any profound effect. Were the level of user charges larger than marginal, we expect the effect to have been significant and in the desired direction. Increased private participation in healthcare expenses however, assumes that people would secure themselves for their healthcare expenses with some precautionary liquid savings. In Chapter 3 we showed that health status is a positive and significant determinant of liquid pre-retirement savings which suggests that individuals save more if they are healthy. It implies that **the government should prioritize prevention if private participation in healthcare expenses increases.** Even though we analyzed a pre-retirement segment of population exclusively, among which we expect the strongest effect because they anticipate a period without earned income, the conclusions may be extended to other cohorts too. If people do not have precautionary liquid savings for private healthcare expenses, any health disorder may put them into poverty or may cause a decrease in their health status because of foregoing care.

Appendix

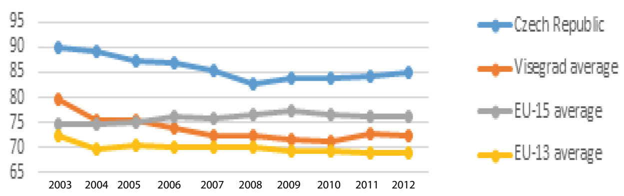
Stylized facts

In 2012 as much as 11,346.31 mil EUR/ 16,720.57 EUR in PPS (EUROSTAT, 2003–2017) was spent on healthcare in the Czech Republic. It makes 1,591.61 EUR per capita in PPS (EUROSTAT, 2003–2017). In relative terms, per capita expenditures on healthcare in 2012 was 20 % above the average of the Visegrad countries (including Czech Republic, Poland, Hungary and Slovakia), however, 43 % below EU-15 average (OECD, 2000–2012). Over the last 13 years, the per capita expenditures on healthcare in the Czech Republic has followed an increasing trend of other EU countries, however.

When expressed as a percentage of GDP, the Czech Republic spent 7.41 % of GDP on healthcare in 2010 (OECD, 2000–2012). WHO (2003–2012) estimated healthcare expenditures to be 7.48 and 7.66 % of GDP in 2011 and 2012, respectively. This puts the Czech Republic below the average of EU-28 and EU-15 countries, however on the average of EU-13 and Visegrad countries. Currently, the expenditure on GDP hardly reaches 7 % of GDP. Currently, there are calls particularly from labor unions, to increase the share to at least to 9 %.

Of all the EU countries, the Czech Republic finances most of its healthcare publicly - in 2012 nearly 85 % of total health expenditure was public compared to 76 % in EU-15 countries or 72 % in Visegrad countries (WHO, 2003–2012). As Figure D1 shows, the share of public spending decreased slightly in 2008 when user charges for an ambulatory doctor visits and an inpatient day were introduced, however, afterwards the share of public resources devoted to healthcare increased again once user charges were abolished.

Figure D1. Public-sector health expenditure % of total health expenditure, 2003–2012



Source:WHO (2003–2012), estimates

Of all the private expenditure as much as 93 % was formed by household-out-of-pocket payments in 2012 which is the highest share compared to all EU-averages (WHO, 2003–2012)

As much as 32.6 % of the entire healthcare resources was directed to inpatient care in 2011. Inpatient care was from 96.9 % financed publicly, which results in 31.59 % of all healthcare expenditures being public and directed to inpatient care (WHO, 2003–2012). The Czech Republic copies the general trend of the Visegrad average, however being significantly above it in all respects, i.e. larger share of resources is devoted to inpatient care and out of the amount devoted to inpatient care a larger share is financed publicly.

Comparing inpatient expenditure in 2003 and 2011 in the Czech Republic, a decrease is remarkable. It is connected with a goal to decrease the number of acute care hospital beds and increase the number of long–term beds. These attempts resulted in 12.6 % of 58,832 beds available in the Czech Republic in 2012 being dedicated specifically to long-term patients (UZIS, 2013).

International comparison of the number of hospital beds per 100 000 inhabitants point to the Czech trend to decrease the number of acute beds. Compared to EU–averages however, the Czech Republic still reports the highest numbers. The number of long–term care beds has been increasing since 2010 (in line with the policy target) after a drop in 2008–2009. **The number of Czech long–term care beds per 100 000 inhabitants**, is nevertheless still **significantly below EU–15** but above Visegrad and EU–13 averages (WHO, 2003–2012). Demand for long–term care stemming from aging population is expected to increase the number of long–term care beds in the future, were it not for an increase of other alternatives of care for the elderly (community–based services, home care, etc.).

Inpatient discharges per 100 inhabitants reached 20.5 in 2011 having experienced a steady decrease over 9 years (WHO, 2003–2012). However, it is **still the highest value compared to other EU countries.**

Between 2009–2011, Czech occupancy rate decreased in absolute terms and roughly oscillated around the EU average in relative terms. The Czech average length of stay, copied a decreasing trend of other EU countries over the whole period, however reporting significantly higher numbers. Keeping in mind the continuing policy target to increase efficiency of hospitals in the Czech Republic, **one would expect an increase and a sharper drop in occupancy rate and the acute length of stay, respectively**.

Capital equipment in hospitals, particularly acute hospitals, is aging and is underutilized which also points to a scope for further efficiency improvements of the system. In the period 2003–2012, capital formation expenditure of Czech hospitals was below the EU–old–member–state average but slightly above the EU–new–member state average. Czech general hospitals copy the trend for all Czech hospitals both in absolute values and as a percentage of GDP. Values and percentages of capital formation expenditures of all hospitals and general hospitals separately suggest that **general hospitals drive hospital capital investment** in the Czech Republic. Investment in hospitals of other types is only marginal. Expressed as percentage of GDP, capital investment has been significantly decreasing since 2009.

The Czech healthcare sector as a whole is better equipped with medical technology than the average of Visegrad countries, but fairly worse than

OECD averages (OECD, 2013). In 2011, there were 6.9 Magnetic Resonance Imaging (MRI) units and 14.8 CT scanners per million inhabitants in the Czech healthcare sector, whereas OECD 28, respectively OECD 29 countries report 13.3 MRI units and 23.6 CT. An average Visegrad country reported only 5.4 MRI units and 12.7 CT scanner per million population in 2013.

Data on utilization of the MRI units and CT scanner show that the number of MRI exams per capita in the Czech Republic is lower (39.0 MRI exams and 89.5 CT exams per 1000 population) than average of the OECD 21, respectively 20 OECD countries (55.4 MRI exams and 131.0 CT exams per 1000 population) but by contrast higher than average of the Visegrad countries (31.0 MRI exams and 79.6 CT exams per 1000 population). Only Slovakia reports comparable data to the Czech Republic. Hungary and Poland indicate significantly lower utilization of medical diagnostic devices. On one hand, advanced imaging represents enormous benefits both to providers and patients due to its non-invasive nature, however they may potentially be connected with adverse effects due to increased exposure to radiation. Although average costs of the technology is the lowest upon maximum utilization, i.e. it is used utmost efficiently from the economic perspective, overuse of advanced imaging technologies has been identified as a key driver of healthcare spending in the world (Kumar, 2011), thus particularly developing countries should search for cheaper alternatives. A meta analysis disentangling this relationship however suggest that some technologies indeed are extremely costly in terms of variable costs, but some technologies were found cost neutral and some even cost-saving (Sorenson *et al.*, 2013). Whether increasing utilization of the imaging units in the Czech Republic would increase efficiency of the Czech healthcare system thus depends on a careful cost-benefit analysis and a well-performed health technology assessment.

Tables

Table D1. Variable description of Votápková & Šťastná (2013)

Inputs & outputs	
total inpatient costs	all inpatient excluding capital costs
acute days	sum of intenstive, surgery and non-surgery days
nursing days	long-term care days
doctor-bed ratio	number of doctors per available bed - quality indicator
nurse-bed ratio	number of nurses per available bed - quality indicator
Determinants	
size	hospitals divided into 3 groups acc. the number of patients treated: (1) small (below 10,000), (2) medium (10,000-20,000) and (3) big (above 20,000) 33 observations in each group
nonprofit	dummy 1 if public not–for–profit, 0 if either public for–profit (> 50 % public share)[a] or private for–profit
share of the elderly	share of 65+ in the municipality
university	dummy 1 for a faculty hospital, 0 otherwise
population	number of total inhabitants in the municipality
competition	the number of general hospitals in the region

Note: [a] Starting in 2003, many regional hospitals were corporatized and became de facto a private legal entity. However, regional authorities retained more than 50 % of shares of these newly created joint–stock companies (“a.s.”).

Table D2. Variable description of Mastromarco *et al.* (2019)

Inputs & outputs	
total inpatient costs	all inpatient excluding capital costs
acute patients DRG-case mix index adjusted	inpatient admissions (excluding ambulatory care) adjusted for the DRG case–mix index
nursing patients	long–term care admissions
publications	first principle component of the Principle Component Analysis (PCA) of the data from the Web of Science database. Inputs to PCA are: (i) articles, (ii) meeting abstracts, (iii) letters, reviews, proceedings papers weighted by the share of authors affiliated to the hospital
nurse-bed ratio	number of nurses per available bed - quality indicator
Determinants	
Nonprofit	dummy 1 if public not–for–profit, 0 if either public for–profit (> 50 % public share)[a] or private for–profit
Specialization	dummy 1 if a specialized center situated in the hospital list of specialized centers from the Ministry of Health
University	dummy 1 for a faculty hospital, 0 otherwise
Occupancy	percentage of potential capacity used
2009 or 2010	dummy 1 if observed in 2009 or 2010, 0 otherwise
2009 or 2010 × University	cross variable
2009 or 2010 × Nonprofit	cross variable

Note: [a] Starting in 2003, many regional hospitals were corporatized and became de facto a private legal entity. However, regional authorities retained more than 50 % of shares of these newly created joint–stock companies (“a.s.”).

Conclusions

This dissertation thesis is a collection of three empirical health economics essays. A comprehensive discussion essay connects the results of the remaining essays and the early research of the author, and brings policy recommendations.

The dissertation thesis has a clear policy motivation arising from the reform attempts of the Czech government which address the recommendations of the OECD reports (OECD, 2011, 2014). The dissertation thesis analytically assesses the effect of the reforms and the areas of the Czech healthcare system which still deserve to be improved.

The first essay in Chapter 1 assesses efficiency of Czech general hospitals using the conditional FDH approach. It is a follow-up research which started already during my master studies and was published as Votápková & Šťastná (2013). The datasets analyzed in the two studies largely overlap but the methodology differs and so does a set of variables. In the latter the variables were enhanced particularly for the DRG case-mix index and publication outputs. The DRG-case mix index, even though not yet optimally designed, and the publication output significantly improved the measured efficiency of specialized, university and big hospitals. However, big and teaching hospitals would still benefit from a different funding scheme.

Chapter 2 assesses the effect of inpatient user charges on the number of patient days of hospitals. We found that user charges exerted only a small and statistically unstable effect on the number of inpatient days that hospitals report. Consistent with other research on Czech inpatient user charges in different sectors of healthcare, we assume that the user charge for an inpatient day was too low to reveal any profound and statistically significant results.

Chapter 3 found that health status is a positive determinant of pre-retirement liquid saving. It also pointed out to the important role of prevention within the healthcare system. We applied a 2SLS methodology, because alternative endogeneity tests revealed that health status is endogenous to savings.

If people are healthy, they more likely save, particularly shortly prior to retirement, to secure themselves for the years when they are not able to work anymore. Likewise, if not healthy, the individuals would prefer spending money to increase their health status at the expense of savings. Effective preventive programs could therefore help ease the tight public healthcare and pension budgets by increasing personal liquid savings. Even though the effect is expected to be the strongest among the pre-retirement segment of the population, effective preventive programs could also boost savings of the remaining cohorts.

Chapter 4 discusses and connects the results of the preceding essays providing also a comprehensive policy recommendations resulting from dissertation thesis and our early work.

There is evidence that the Czech healthcare sector lacks behind other EU countries in some major characteristics, including the low share of GDP devoted to healthcare, low private participation on healthcare expenses, excessive number of patient days, etc. OECD (2018) highlights the need for increased financial sustainability of the Czech healthcare system. To reach it, the OECD (2018) recommends, among other things, to (1) address the challenge of aging population; (2) improve hospital management; (3) reimburse hospitals based on performance rather than on a flat rate, (4) strengthen the role of general practitioners in prevention and care for the chronically ill; and (5) equalize availability and quality of care across regions.

Financial sustainability as the ultimate recommendation of OECD (2018) penetrates the whole dissertation thesis. Each chapter of this dissertation thesis analyzes some of the problems highlighted by OECD (2018) based on particular reforms or situations.

Specifically, in Chapter 1 and Section 4.1, we show that the DRG case-mix index is crucial as a performance measure. It is clear that once employed as a reimbursement mechanism, efficiency of the Czech healthcare system will increase. Section 4.1 highlights the need for the DRG as the reimbursement mechanism. We also address its benefits for the efficiency of capital investments.

Our research both within this thesis and the early research of the author brings quantitative evidence of the benefits of private participation in healthcare expenses for financial sustainability of the system. Chapter 3 also quantitatively showed the importance of prevention for the economic growth of a country.

Other research of the author which is beyond the scope of the thesis addresses some of the remaining recommendations of the OECD (2018). Chronic illnesses, such as Type II Diabetes Mellitus, is analyzed in Votápková *et al.* (2017) stressing the need for a reform of care for chronically ill, otherwise chronic diseases will represent a large economic burden for the public budget in the near future.

Once published, essays included in this dissertation thesis will be disseminated among policy-makers to contribute to health policy discussions and generate similar policy impact as our previous work which was cited by OECD (2018) report, namely Votápková & Žílová (2012, 2016b).

References

Adams, P., M. D. Hurd, D. McFadden, A. Merrill, & T. Ribeiro (2003): "Healthy, wealthy, and wise? tests for direct causal paths between health and socioeconomic status." *Journal of Econometrics* **112(1)**: pp. 3–56.

Aiken, L. H., S. P. Clarke, D. M. Sloane, J. Sochalski, & J. H. Silber (2002): "Hospital nurse staffing and patient mortality, nurse burnout and job dissatisfaction." *JAMA* **288(16)**: pp. 1987–1999.

Aitchison, J. & C. G. G. Aitken (1976): "Multivariate binary discrimination by the kernel method." *Biometrika* **63(3)**: pp. 413–420.

Alexa, J., L. Rečka, J. Votápková, E. van Ginneken, A. Spranger, & F. Wittenbecher (2015): "Czech republic: Health system review." *Health Systems in Transition* **17(1)**: pp. 1–165.

Anthun, K. S., S. A. C. Kittelsen, & J. Magnussen (2017): "Productivity growth, case mix and optimal size of hospitals. A 16-year study of the norwegian hospital sector." *Health Policy* **121(4)**: pp. 418–425.

Araújo, C., C. P. Barros, & P. Wanke (2014): "Efficiency determinants and capacity issues in Brazilian for-profit hospitals." *Health Care Manag Sci* **17(2)**: pp. 126–138.

Ashenfelter, O. & D. Card (1985): "Using the longitudinal structure of earnings to estimate the effect of training programs." *The Review of Economics and Statistics* **67(4)**: pp. 648–60.

Atella, V., M. Brunetti, & N. Maestas (2012): "Household portfolio choices, health status and health care systems: A cross-country analysis based on {SHARE}." *Journal of Banking & Finance* **36(5)**: pp. 1320 – 1335.

Augurzky, B., T. Bauer, & S. Schaffner (2006): "Copayments in the German health system: Does it work?" *RWI Discussion Papers Series* **43**.

Bates, L. J., K. Mukherjee, & E. Santerre, Rexford (2006): "Market structure and technical efficiency in the hospital services industry: A dea approach." *Med Care Res Rev* **63(4)**: pp. 499–524.

BECK, N., F. HOEIJMAKERS, E. M. VAN DER WILLIK, D. J. HEINEMAN, J. BRAUN, R. A. TOLLENAAR, W. H. SCHREURS, & M. W. WOUTERS (2018): "National comparison of hospital performances in lung cancer surgery: The role of case mix adjustment." *Ann Thorac Surg* **106(2)**: pp. 412–420.

BECK, R. G. & J. M. HORNE (1980): "Utilization of publicly insured health services in saskatchewan before, during and after copayments." *Medical Care* **18**: pp. 787–806.

BLANK, J. L. T. & V. VALDMANIS (2010): "Environmental factors and productivity on dutch hospitals: a semi-parametric approach." *Health Care Management Science* **13**: pp. 27–34.

BLUNDELL, R., R. CRAWFORD, E. FRENCH, & G. TETLOW (2016a): "Comparing retirement wealth trajectories on both sides of the pond." *Fiscal Studies* **37(1)**: pp. 105–130.

BLUNDELL, R., R. CRAWFORD, E. FRENCH, & G. TETLOW (2016b): "Comparing retirement wealth trajectories on both sides of the pond." *Fiscal Studies* **37(1)**: pp. 105–130.

BLUNDELL, R. & M. C. DIAS (2008): "Alternative approaches to evaluation in empirical microeconomics." *CEF.UP Working Papers 0805*, Universidade do Porto, Faculdade de Economia do Porto.

BORCHERS, A. & M. GERSHWIN (2012): "Sociological differences between women and men: Implications for autoimmunity." *Autoimmunity Reviews* **11(6)**: pp. A413–A421.

BRUNING, E. R. & C. A. REGISTER (1989): "Technical efficiency within hospitals: Do profit incentives matter?" *Appl Econ* **21(9)**: pp. 1217–1233.

BĂDIN, L., C. DARAIO, & L. SIMAR (2010): "Optimal bandwidth selection for conditional efficiency measures: A data driven approach." *Eur J Oper Res* **201**: pp. 633–40.

BĂDIN, L., C. DARAIO, & L. SIMAR (2012): "How to measure the impact of environmental factors in a nonparametric production model?" *Eur J Oper Res* **223**: pp. 818–833.

BĂDIN, L., C. DARAIO, & L. SIMAR (2014): "Explaining inefficiency in nonparametric production models: The state of the art." *Annal Oper Res* **1**: pp. 5–30.

BURGESS, J. F. J. & P. WILSON (1995): "Decomposing hospital productivity changes, 1985-1988: A nonparametric Malmquist approach." *J Prod Anal* **6(4)**: pp. 343–363.

BÖRSCH-SUPAN, A. (2016): "Survey of health, ageing and retirement in europe (share) wave 5. release version: 5.0.0. share-eric. data set." *Technical report*, SHARE.

Börsch-Supan, A. & A. Lusardi (2003): "Chapter 1 - saving: A cross-national perspective." In A. Börsch-Supan (editor), "Life-Cycle Savings and Public Policy," pp. 1 – 31. Boston: Academic Press.

Cazals, C., J.-P. Florens, & L. Simar (2002): "Nonparametric Frontier Estimation: a Robust Approach." *Journal of Econometrics* **106(1)**: pp. 1–25.

Cheng, S.-H., C.-C. Chen, & S.-L. Tsai (2012): "The impacts of drg-based payments on health care provider behaviors under a universal coverage system: a population-based study." *Health policy* **107(2)**: pp. 202–208.

Cherkin, D., L.Grothaus, & E. Wagner (1990): "The effect of office visit copayments on preventive care services in an hmo." *Inquiry* **27**: pp. 24–38.

Chirikos, T. N. & A. M. Sear (2000): "Measuring Hospital Efficiency: A Comparison of Two Approaches." *Health Service Research* **34:6**: pp. 1389–1408.

Choi, J. H., S. M. Fortsch, I. Park, & I. Jung (2017): "Efficiency of U.S. hospitals between 2001 and 2011." *Manag Decis Econ* **38:8**: pp. 1071–1081.

Chowdhury, H. & V. Zelenyuk (2016): "Performance of hospital services in Ontario: DEA with truncated regression approach." *Omega* **63**: pp. 111–122.

Chowdhury, H., V. Zelenyuk, A. Laporte, & W. P. Wodchis (2014): "Analysis of productivity, efficiency and technological changes in hospital services in Ontario: How does case-mix matter?" *Int J Prod Econ* **150**: pp. 74–82.

Clement, J., V. Valdmanis, G. Bazzoli, M. Zhao, & A. Chukmaitov (2008): "Is more better? an analysis of hospital outcomes and efficiency with a dea model of output congestion." *Health Care Management Science* **11(1)**: pp. 67–77.

Coile, C. & K. Milligan (2006): "How household portfolios evolve after retirement: The effect of aging and health shocks." *Working Paper 12391*, National Bureau of Economic Research.

Constatinides, G. M., M. Harris, & R. M. Stulz (2012): *Handbook of the Economics of Finance, Volume 2A*. North Holland.

Cordero, J. M., E. Alonso-Morán, R. Nuño Solinis, J. F. Orueta, & R. S. Arce (2015): "Efficiency assessment of primary care providers: A conditional nonparametric approach." *Eur J Oper Res* **240(1)**: pp. 235–244.

Curtis, C. C., S. Lugauer, & N. C. Mark (2015): "Demographic patterns and household saving in china." *American Economic Journal: Macroeconomics* **7(2)**: pp. 58–94.

Czypionka, T., M. Kraus, S. Mayer, & G. Röhrling (2014): "Efficiency, ownership, and financing of hospitals: The case of Austria." *Health Care Manag Sci* **17(4)**: pp. 331–347.

Daidone, S. & F. D'Amico (2009): "Technical efficiency, specialization and ownership form: Evidences from a pooling of Italian hospitals." *J Prod Anal* **32(3)**: pp. 203–216.

Daouia, A. & I. Gijbels (2011): "Robustness and inference in nonparametric partial-frontier modeling." *J Econom* **161**: pp. 147–165.

Daouia, A. & L. Simar (2007): "Nonparametric efficiency analysis: A multivariate conditional quantile approach." *J Econom* **140**: pp. 375–400.

Daraio, C. & L. Simar (2005): "Introducing environmental variables in nonparametric frontier models: A probabilistic approach." *J Prod Anal* **24(1)**: pp. 93–121.

Daraio, C. & L. Simar (2007): *Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications*. Studies in productivity and efficiency. Springer.

Daraio, C. & L. Simar (2014): "Directional distances and their robust versions: Computational and testing issues." *Eur J Oper Res* **237**: pp. 358–369.

DeWitte, K. & M. Kortelainen (2013): "What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables." *Appl Econ* **45(17)**: pp. 2401–2412.

Dlouhý, M., J. Jablonský, & I. Novosádová (2007): "Využití analýzy obalu dat pro hodnocení efektivnosti českých nemocnic." *Polit ekon* **1**: pp. 60–71.

Dormont, B. & C. Milcent (2012): "Ownership and Hospital Productivity." *CEPREMAP Working Papers (Docweb) 1205*, CEPREMAP.

Elofsson, S., A.-L. Undén, & I. Krakau (1998): "Patient charges – a hindrance to financially and psychosocially disadvantage groups seeking care." *Social Science & Medicine* **46(10)**: pp. 1375–1380.

EUROSTAT (2003–2017): "Statistics database." http://epp.eurostat.ec.europa.eu/.

Farbmacher, H. (2009): "Copayments for doctor visits in germany and the probability of visiting a physician - evidence from a natural experiment." *Discussion papers in economics*, University of Munich, Department of Economics.

Farsi, M. & M. Filippini (2006): "An analysis of efficiency and productivity in Swiss hospitals." *Swiss J Econ Stat* **142**: pp. 1–37.

Feinstein, J. & D. McFadden (1989): *The Dynamics of Housing Demand by the Elderly: Wealth, Cash Flow, and Demographic Effects*, pp. 55–92. University of Chicago Press.

Feldstein, P. (1971): "Hospital cost inflation: A study of nonprofit price dynamics." *American Economic Review* **61**: pp. 853–72.

Ferreira, D., R. Marques, & A. Nunes (2018): "Economies of scope in the health sector: The case of Portuguese hospitals." *Eur J Oper Res* **266(2)**: pp. 716–735.

Ferreira, D. C. & R. C. Marques (2016): "Should inpatients be adjusted by their complexity and severity for efficiency assessment? Evidence from portugal." *Health Care Manag Sci* **19(1)**: p. 43—57.

Fioroni, T. (2010): "Optimal savings and health spending over the life cycle." *The European Journal of Health Economics* **11(4)**: pp. 355–365.

Fried, H. O., K. C. A. Lovell, & S. S. Schmidt (2008): *The Measurement of Productive Efficiency and Productivity Growth.* Oxford University Press, Inc.

Frohloff, A. (2007): "Cost and Technical Efficiency of German Hospitals: A Stochastic Frontier Analysis." *Ruhr Economic Papers* .

Fu, T.-T., J.-T. Liu, & J. K. Hammitt (1999): "Consumer willingness to pay for low-pesticide fresh produce in taiwan." *Journal of Agricultural Economics* **50(2)**: pp. 220–233.

Gemmill, M. C., S. Thomson, & E. Mossialos (2008): "What impact do prescription drug charges have on efficiency and equity? evidence from high-income countries." *International Journal for Equity in Health* **7**: p. 12.

Greene, W. (2002): *Econometric Analysis (5th edition).* Prentice Hall.

Grolnick, W. S., C. Benjet, C. O. Kurowski, & N. H. Apostoleris (1997): "Predictors of parent involvement in children' s schooling." *Journal of Educational Psychology* **89**: pp. 538–548.

Grosskopf, S., D. Margaritis, & V. Valdmanis (2001a): "Comparing teaching and non-teaching hospitals: A frontier approach (teaching vs. non-teaching hospitals)." *Health Care Manag Sci* **4(2)**: pp. 83–90.

Grosskopf, S., D. Margaritis, & V. Valdmanis (2001b): "The effects of teaching on hospital productivity." *Socioecon Plan Sci* **35(3)**: pp. 189–204.

Grosskopf, S. & V. Valdmanis (1993): "Evaluating hospital performance with case-mix-adjusted outputs." *Med Care* **31(6)**: pp. 525–532.

GRUBER, J. & H. KAISER (2006): "The role of consumer copayments for health care. lessons from the rand health insurance experiment and beyond." *Henry J. Kaiser Family Foundation* .

GUISO, L., M. HALIASSOS, & T. JAPPELLI (2002): *Household portfolios.* Cambridge MA: MIT Press.

HALKOS, G. E. & N. G. TZEREMES (2011): "A conditional nonparametric analysis for measuring the efficiency of regional public healthcare delivery: An application to greek prefectures." *Health Policy* **103(1)**: pp. 73 – 82.

HALL, P., J. RACINE, & Q. LI (2004): "Cross-validation and the estimation of conditional probability densities." *J Am Stat Assoc* **99(468)**: pp. 1015–1026.

HARWELL, M. R. & G. G. GATTI (2001): "Rescaling ordinal data to interval data in educational research." *Review of Educational Research* **71**: pp. 538–548.

HELMS, J., J. NEWHOUSE, & C. E. PHELPS (1978): "Copayments and demand for medical care: The california medicaid experience." *Bell Journal of Economics* **9(1)**: pp. 192–208.

HOFMARCHER, M. M., I. PETERSON, & M. RIEDEL (2002): "Measuring Hospital Efficiency in Austria - A DEA Approach." *Health Care Management Science* **5**: pp. 7–14.

HOLLINGSWORTH, B. (2008): "The measurement of efficiency and productivity of health care delivery." *Health Econ* **17(10)**: pp. 1107–1128.

HROMÁDKOVÁ, E. (2016): *Essays on Cost-Containment Measures in Health Care*. Ph.D. thesis.

INSTITUTE OF HEALTH INFORMATION AND STATISTICS OF THE CZECH REPUBLIC (2011): "Zdravotnictví České republiky 2010 ve statistických údajích." *Technical report*, ÚZIS ČR.

INSTITUTE OF HEALTH INFORMATION AND STATISTICS OF THE CZECH REPUBLIC, REGIONAL OFFICES (2006-2010): "Zdravotnictví (kraje + ČR)." *Technical report*, ÚZIS ČR.

JACOBS, R. (2001): "Alternative methods to examine hospital efficiency: Data envelopment analysis and stochastic frontier analysis." *Health Care Manag Sci* **4**: pp. 103–116.

JACOBS, R. & D. DAWSON (2003): "Hospital efficiency targets." *Health Econ* **12(8)**: pp. 669–684.

JAMIESON, S. (2004): "Likert scales: How to (ab)use them." *Medical Education* **38**: pp. 1212–1218.

Janlöv, N. (2007): "Performance in Swedish health care - Quantity versus quality." *iHEA 2007 6th World Congress: Explorations in Health Economics* .

Jeong, S., B. Park, & L. Simar (2010): "Nonparametric conditional efficiency measures: Asymptotic properties." *Ann Oper Res* **173**: pp. 105–122.

Junji, K. (2003): "The effects of a continuous increase in lifetime on saving." *Review of Income and Wealth* **49(2)**: pp. 163–183.

Kalousová, L. (2014): "Curing over-use by prescribing fees: An evaluation of the effect of user fees' implementation on healthcare use in the Czech Republic." *Health policy and planning* pp. 1–9.

Keeler, E. B. & J. E. Rolph (1988): "The demand for episodes of treatment in the health insurance experiment." *Journal of Health Economics* **7(4)**: pp. 337 – 367.

Kelejian, H. H. (1971): "Two-stage least squares and econometric systems linear in parameters but nonlinear in the endogenous variables." *Journal of the American Statistical Association* **66(334)**: pp. 373–374.

Knolll, M., R. Tamborini, Christopher, & K. Whitman (2012): "I do...want to save: Marriage and retirement savings in young households." *Journal of Marriage and Family* **74**: pp. 86–100.

Kohl, S., J. Schoenfelder, A. Fügener, & J. O. Brunner (2018): "The use of Data Envelopment Analysis (DEA) in healthcare with a focus on hospitals." *Health Care Manag Sci* .

Kooreman, P. (1994): "Nursing Homecare in The Netherlands: a Nonparametric Efficiency Analysis." *Journal of Health Economics* **13**: pp. 301–316.

Kožený, P., J. Němec, J. Kárníková, & M. Lomíček (2010): *Klasifičkní systém DRG*. Grada Publishing, a.s.

Kumar, R. K. (2011): "Technology and healthcare costs." *Annals of Pediatric Cardiology* **4(1)**: pp. 84–86.

Kuzon, W. M., M. G. Urbanchek, & S. McCabe (1996): "The seven deadly sins of statistical analysis." *Annals of Plastic Surgery* **37**: pp. 265–272.

Lee, M. L. (1971): "A conspicuous production theory of hospital behavior." *Southern Economic Journal* **38(1)**: pp. 48–58.

Li, D., L. Simar, & V. Zelenyuk (2016): "Generalized nonparametric smoothing with mixed discrete and continuous data." *Comput Stat Data Anal* **100**: pp. 424–444.

Li, Q. & J. Racine (2004): "Cross-validated local linear nonparametric regression." *Stat Sin* **14(3)**: pp. 485–510.

Li, Q. & J. S. Racine (2007): *Nonparametric econometrics: Theory and practice*. Princeton University Press, Princeton.

Li, Q. & J. S. Racine (2008): "Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data." *J Bus Econ Stat* **26(4)**: pp. 423–434.

Linna, M. (1998): "Measuring hospital cost efficiency with panel data models." *Health Econ* **7(5)**: pp. 415–427.

Linna, M., U. Häkkinen, M. Peltola, J. Magnussen, K. S. Anthun, S. Kittelsen, A. Roed, K. Olsen, E. Medin, & C. Rehnberg (2010): "Measuring cost efficiency in the Nordic hospitals—a cross-sectional comparison of public hospitals in 2002." *Health Care Manag Sci* **13(4)**: pp. 346–357.

Linna, M. & U. Häkkinen (1998): "A comparative application of econometric frontier and DEA methods for assessing cost efficiency of finnish hospitals." *Dev Health Econ Public Policy* **6**: pp. 169–187.

Linna, M., U. Häkkinen, & E. Linnakko (1998): "An econometric study of costs of teaching and research in Finnish hospitals." *Health Econ* **7(4)**: pp. 291–305.

Louis, D. Z., E. J. Yuen, M. Braga, A. Cicchetti, C. Rabinowitz, C. Laine, & J. S. Gonnella (1999): "Impact of a drg-based hospital financing system on quality and outcomes of care in italy." *Health services research* **34(1 Pt 2)**: p. 405.

Magnussen, J. (1996): "Efficiency measurement and the operationalization of hospital production." *Health Serv Res* **31:1**: pp. 21–37.

Mana, M. & V. Kalnicka (2016): "Vysledky zdravotnickych uctu CR v letech 2010 az 2014." *Technical report*, Czech statistical office.

Manning, Willard G, e. a. (1987): "Health insurance and the demand for medical care: Evidence from a randomized experiment." *American Economic Review* **77(3)**: pp. 251–77.

Mastromarco, C. & L. Simar (2015): "Effect of FDI and time on catching-up: New insights from a conditional nonparametric frontier analysis." *J Appl Econom* **30**: pp. 826–847.

Mastromarco, C., L. Šťastná, & J. Votápková (2019): "Efficiency of hospitals in the czech republic: Conditional efficiency approach." *Journal of Productivity Analysis* **51(1)**: pp. 73–89.

Matranga, D., F. Bono, A. Casuccio, A. Firenze, L. Marsala, R. Giaimo, F. Sapienza, & F. Vitale (2014): "Evaluating the effect of organization and context on technical efficiency: A second-stage DEA analysis of Italian hospitals." *Epidemiol Biostat Public Health* **11(1)**.

Meer, J., D. L. Miller, & H. S. Rosen (2003): "Exploring the health-wealth nexus." *Working Paper 9554*, National Bureau of Economic Research.

Merton, R. C. (1969): "Lifetime portfolio selection under uncertainty: The continuous-time case." *The Review of Economics and Statistics* **51(3)**: pp. 247–257.

Michaud, P.-C. & A. van Soest (2004): "Health and wealth of elderly couples: Causality tests using dynamic panel data models." *IZA Discussion Papers 1312*, Institute for the Study of Labor (IZA).

Milligan, K. S. (2005): "Life-cycle asset accumulation and allocation in canada." *Canadian Journal of Economics* **38(3)**: pp. 1057–1106.

Nakajima, M. & I. A. Telyukova (2020): "Home equity in retirement." *International Economic Review* .

Nayar, P. & Y. A. Ozcan (2008): "Data Envelopment Analysis Comparison of Hospital Efficiency and Quality." *J. Med. Syst.* **32(3)**: pp. 193–199.

Needleman, J., P. Buerhaus, S. Pankratz, C. L. Leibson, S. R. Stevens, & M. Harris (2011): "Nurse staffing and inpatient hospital mortality." *The New England Journal of Medicine* **364**: pp. 1037–1045.

Newhouse, J. P. (1970): "Toward a theory of nonprofit institutions: An economic model of a hospital." *The American Economic Review* **60(1)**: pp. 64–74.

Newhouse, J. P. & R. C. I. E. Group (1993): *Free for all?: Lessons from the RAND health insurance experiment.* Harvard Univ Pr.

Novosádová, I. & M. Dlouhý (2007): "Hodnotenie technickej efektívnosti nemocníc a odmeňovanie zdravotníkov." *Ekon cas* **55 (8)**: pp. 762–782.

O'Brien, B. (1989): "The effect of patient charges on the utilisation of prescription medicines." *Journal of Health Economics* **8(1)**: pp. 109–132.

OECD (2000–2012): "OECD.StatExtracts." http://www.oecd-ilibrary.org/social-issues-migration-health/data/oecd-health-statistics_health-data-en.

OECD (2011): "Enhancing public spending efficiency." In "OECD, Economic Surveys: Czech Republic 2011," OECD.

OECD (2013): "Health at a Glance 2013." *Technical report*, OECD Health Statistics.

OECD (2014): "OECD Reviews of Healthcare Quality: Czech Republic 2014: Raising Standards." *Technical report*, OECD.

OECD (2018): "OECD Economic Surveys: Czech Republic." *Technical report*, OECD.

Pang, G. & M. Warshawsky (2010): "Optimizing the equity-bond-annuity portfolio in retirement: The impact of uncertain health expenses." *Insurance: Mathematics and Economics* **46(1)**: pp. 198–209.

Park, B., L. Simar, & C. Weiner (2000): "The FDH estimator for productivity efficiency scores: Asymptotic properties." *Econom theory* **16**: pp. 855–877.

Pauly, M. & M. Redisch (1973): "The not-for-profit hospital as a physicians' cooperative." *The American Economic Review* **63(1)**: pp. 87–99.

Pearson, K. (1901): "On lines and planes of closest fit to systems of points in space." *Philos Mag* **2**: pp. 559–572.

Poterba, J., S. Venti, & D. Wise (2011): "The composition and drawdown of wealth in retiremen." *Journal of Economic Perspectives* **25(4)**: pp. 95–118.

Poterba, J. M. & A. A. Samwick (1997): "Household portfolio allocation over the life cycle." *Working Paper 6185*, National Bureau of Economic Research.

R Development Core Team (2006): *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Racine, J. (1997): "Consistent significance testing for nonparametric regression." *J Bus Econ Stat* **15(3)**: pp. 369–378.

Racine, J. & Q. Li (2004): "Nonparametric estimation of regression functions with both categorical and continuous data." *J Econom* **119(1)**: pp. 99–130.

Racine, J. S., J. Hart, & Q. Li (2006): "Testing the significance of categorical predictor variables in nonparametric regression models." *Econom Rev* **25(4)**: pp. 523–544.

Reder, M. (1965): "Some problems in the economics of hospitals." *American Economic Review* **55**: pp. 472–80.

Ricketts, C. F., J. P. Rezek, & R. C. Campbell (2013): "The influence of individual health outcomes on individual savings behavior." *The Social Science Journal* **50**: pp. –.

Roemer, M. I., C. E. Hopkins, L. Carr, & F. Gartside (1975): "Copayments for ambulatory care: Penny-wise and pound-foolish." *Medical Care* **13**: pp. 457–466.

ROSEN, H. S. & S. WU (2004): "Portfolio choice and health status." *Journal of Financial Economics* **72(3)**: pp. 457–484.

ROSENBERG, M. A. & M. J. BROWNE (2001): "The impact of the inpatient prospective payment system and diagnosis-related groups: a survey of the literature." *North American Actuarial Journal* **5(4)**: pp. 84–94.

ROSKO, M. D. (2001): "Cost efficiency of US hospitals: A stochastic frontier approach." *Health Econ* **10(6)**: pp. 539–551.

ROSKO, M. D. & J. A. CHILINGERIAN (1999): "Estimating Hospital Inefficiency: Does Case Mix Matter?" *J Med Syst* **23(1)**: pp. 57–71.

RYAN, A. M., J. F. BURGESS, & J. DIMICK (2015): "Why we should not be indifferent to specification choices for difference-in-differences." *Health services research* **50(4)**.

SALM, M. (2010): "Subjective mortality expectations and consumption and saving behaviours among the elderly." *Canadian Journal of Economics* **43(3)**: pp. 1040–1057.

SALTMAN, R. B. & J. FIGUERAS (1997): "European health care reform : analysis of current strategies." *Technical report*, Observatory for Health Systems and Policies.

SAMUELSON, P. (1969): "Lifetime portfolio selection by dynamic stochastic programming." *The Review of Economics and Statistics* **51(3)**: pp. 239–46.

SCHREYÖGG, J. & M. M. GRABKA (2008): "Copayments for ambulatory care in germany: A natural experiment using a difference-in-difference approach." *Discussion Papers of DIW Berlin 777*, DIW Berlin, German Institute for Economic Research.

SCITOVSKY, A. & N. MCCALL (1977): "Coinsurance and the demand for physician services: four years later." *Soc Secur Bull* **40(5)**: pp. 19–27.

SHESHINSKI, E. (2009): "Longevity and aggregate savings." *Discussion Paper Series dp519*, The Center for the Study of Rationality, Hebrew University, Jerusalem.

SIMAR, L. & P. W. WILSON (2007): "Estimation and inference in two-stage, semi-parametric models of production processes." *J Econom* **136(1)**: pp. 31–64.

SMITH, J. P. (1999): "Healthy bodies and thick wallets: The dual relation between health and economic status." *Journal of Economic Perspectives* **13(2)**: pp. 145–166.

SORENSON, C., M. DRUMMOND, & B. B. KHAN (2013): "Medical technology as a key driver of rising health expenditure: disentangling the relationship." *ClinicoEconomics and Outcomes Research* **5**: pp. 223–234.

STEVENS, S. S. (1946): "On theory of scales of measurement." *Science* **103**: pp. 677–680.

TAMBOR, M., M. PAVLOVA, P. WOCH, & W. GROOT (2010): "Diversity and dynamics of patient cost-sharing for physicians' and hospital services in the 27 european union countries." *The European Journal of Public Health* .

TIEMANN, O. & J. SCHREYÖGG (2012): "Changes in hospital efficiency after privatization." *Health Care Manag Sci* **15(4)**: pp. 310–326.

TRIVEDI, A. N., H. MOLOO, & V. MOR (2010): "Increased ambulatory care copayments and hospitalizations among the elderly." *New England Journal of Medicine* **362(4)**: pp. 320–328.

VALDMANIS, V. (1992): "Sensitivity analysis for DEA models: An empirical example using public vs. NFP hospitals." *J Public Econ* **48**: pp. 185–205.

VAN BUUREN, S., J. P. L. BRANDS, C. G. M. GROOTHUIS-OUDSHOORN, & D. RUBIN (2006): "Fully conditional specification in multivariate imputation." *Journal of Statistical Computation and Simulation* **76(2)**: pp. 1049–1064.

VARABYOVA, Y., C. R. BLANKART, A. TORBICA, & J. SCHREYÖGG (2017): "Comparing the efficiency of hospitals in Italy and Germany: Nonparametric conditional approach based on partial frontier." *Health Care Manag Sci* **20(3)**: pp. 379–394.

VENTI, S. F. & D. A. WISE (1987): "Aging, moving, and housing wealth." *Working Paper 2324*, National Bureau of Economic Research.

VENTI, S. F. & D. A. WISE (1989): "But they don't want to reduce housing equity." *Working Paper 2859*, National Bureau of Economic Research.

VENTI, S. F. & D. A. WISE (2000): "Aging and housing equity." *Working Paper 7882*, National Bureau of Economic Research.

VITALIANO, D. F. & M. TOREN (1996): "Hospital cost and efficiency in a regime of stringent regulation." *East Econ J* **22:2**: pp. 161–175.

VOS, T., J. BARENDREGT, C. MIHALOPOULOS, L. VEERMAN, A. MAGNUS, L. COBIAC, M. BERTRAM, & A. WALLACE (2010): "Assessing cost-effectiveness in prevention : Ace–prevention september 2010 final report." *Technical report*, University of Queensland & Deakin University.

VOTÁPKOVÁ, J. (2019): "The effect of inpatient user charges on inpatient care." *IES WP, submitted* .

VOTÁPKOVÁ, J., P. HROBOŇ, & P. KUČOVÁ (2017): "Pacienti s diabetem mellitem 2. typu v Čr - prevalence stanovená na základě spotřeby léčiv a náklady na tyto pacienty." *Ekonomie ve zdravotnictví* **4**: pp. 7–13.

Votápková, J. & L. Šťastná (2013): "Efficiency of hospitals in the czech republic." *Prague Economic Papers* **4**.

Votápková, J. & P. Žílová (2012): "The effect of the abolition of user charges on the demand for ambulatory doctor visits." *IES Working Papers* .

Votápková, J. & P. Žílová (2015): "Health care efficiency in the czech republic - evidence for inpatient care." *IES Occassional Paper 2/2015* .

Votápková, J. & P. Žílová (2016a): "The abolition of user charges and the demand for ambulatory visits: evidence from the czech republic." *Health Economics Review* **6(1)**: p. 29.

Votápková, J. & P. Žílová (2016b): "Determinants of generic substitution in the czech republic." *IES WP* **5**.

Whitaker, E., J. L. Bokemeiner, & S. Loveridge (2012): "Interactional associations of gender on savings behavior: Showing gender´s continued influence on economic action." *Journal of Family and Economic Issues* pp. 1–15.

WHO (2003–2012): "Health for all database." http://data.euro.who.int/hfadb/.

Wilson, P. W. (1993): "Detecting Outliers in Deterministic Nonparametric Frontier Models with Multiple Outputs." *Journal of Business & Economic Statistics* **11(3)**: pp. 319–23.

Wing, C., K. Simon, & R. A. Bello-Gomez (2018): "Designing difference-in-differences studies: Best practices for public health policy research." *Annual Review of Public Health* **39**: pp. 453–69.

Winkelmann, R. (2003): "Co-payments for prescription drugs and the demand for doctor visits - evidence from a natural experiment." *Working Papers 0307*, University of Zurich, Socioeconomic Institute.

Wooldridge, J. M. (2002): *Econometric analysis of cross section and panel data.* Cambridge and London: MIT Press.

Worthington, A. C. (2004): "Frontier Efficiency Measurement in Health Care: A Review of Empirical Techniques and Selected Applications." *Med Care Res Rev* **61(2)**: pp. 135–170.

Wu, H. & S.-O. Leung (2017): "Can likert scales be treated as interval scales?—a simulation study." *Journal of Social Service Research* **43(4)**: pp. 527–532.

Wu, S. (2001): "The effects of health events on the economic status of married couples." *Journal of Human Resources* **38**.

Yong, K. & A. Harris (1999): "Efficiency of Hospitals in Victoria under Casemix Funding: A Stochastic Frontier Approach." Working paper 92. Centre for Health Program Evaluation.

Zhang, J. (2007): "A did analysis of the impact of health insurance reform in the city of hangzhou." *Health Economics* **16(12)**: pp. 1389–1402.

Zuckerman, S., J. Hadley, & L. Iezzoni (1994): "Measuring hospital efficiency with frontier cost functions." *J Health Econ* **13**: pp. 255–280.

Zweifel, P. & W. G. Manning (2000): "Chapter 8 - moral hazard and consumer incentives in health care* *the authors wish to thank andrew jones, joe newhouse, edward norton, and wynand van de ven for excellent suggestions and criticisms." In A. J. Culyer & J. P. Newhouse (editors), "Handbook of Health Economics," volume 1 of *Handbook of Health Economics*, pp. 409 – 459. Elsevier.

Zápal, J. (2010): "Doctor-Visit Co-Payment Exemption for Children: First Look at the Data." *Czech Journal of Economics and Finance (Finance a uver)* **60(1)**: pp. 58–72.

Response to reviewers

I thank the reviewers for their valuable comments on the pre-defense version of the thesis. The suggestions were incorporated as follows:

Changes to Chapter 1

Response to Olena Stavrunova

- “Page 23. Please check if the scale of the X-axis is correct in the figure. I think the numbers are missing one zero (e.g. it should be not 2000 but 20000).”

Thank you. Corrected

- “Be more careful about interpretation of the finding that for-profit hospitals exhibit higher efficiency. The author seems to attribute this difference in efficiency to the restructuring that took place after the hospitals became for-profit. However, it is also possible that more efficient hospitals chose to become for-profit”

The decision to corporatize was purely political, thus external to the hospital management. Until 2000, most of Czech hospitals were state-owned. After Czech regions ("kraje") were established in January 2000, many of state-owned hospitals were transferred to their ownership. Some hospitals were overtaken by municipalities. “Kraje” and municipalities then decided to corporatize hospitals into joint-stock companies hoping in more transparent accounting and financial stability of the hospitals. If the administrative unit decided to corporatize, it did so for all the hospitals in their hands. However, not all regions or municipalities decided to corporatize. The self-selection bias thus does not hold here. The explanation was added to section 1.2.2.

Response to Andrea Menclová

- “Add references to seminal health economics papers on the objective functions (and hence behavior) of non-profit hospitals. (see report)”

References were added in section 1.2.2

Response to Michal Horný

- “Although such a definition of efficiency is technically correct, the selected output metrics do not capture the main purpose of the health care system - production of health. Hypothetically speaking, a hospital that uses very few resources to provide poor quality care to many patients may be marked as ’efficient’ in terms of the number of patients treated per unit of total inpatient costs, but such scenario is not an aim for which we should strive. I encourage the author to add a brief discussion of various efficiency metrics, their corresponding advantages, and disadvantages, as well as limitations of their interpretation.”

Output of hospitals should ideally reflect increase to patient’s health status. However, it is technically impossible to measure, thus in all hospital efficiency studies, intermediate output is applied instead. Efficiency studies employ either admissions or inpatient days as a measure of inpatient care output. The latter, however, may suffer from endogeneity born by the length of stay. The discussion was added to section 1.2.1.
Quality of hospital output is important too. By definition of the efficiency model, quality usually represents another dimension of outputs thus increasing efficiency of those hospitals which produce less output of higher quality, i.e. units which would be considered as inefficient without this additional output dimension. Quality of care has been accounted for differently in the literature. For instance, Zuckerman *et al.* (1994) used mortality rates, Vitaliano & Toren (1996) applied technology index and occupancy rate, Frohloff (2007) employed a nurse-bed ratio.
Quality indicators included in the analysis cover nurse-bed ratio which is included among outputs since it directly determines it. Others are included among determinants of inefficiency (e.g. specialization).

- “I would like to ask the author to elaborate on the specific reasons why more than half (57 %) of hospitals in the sample were excluded from the analysis. The author provides reasons such as that some hospitals were closed, some did not report data, and some observations were deemed to be outliers. Hospital closure is a valid reason for exclusion. Excluding observations with missing data should not affect the result and conclusions if the data were missing at random. Finally, outliers are actual observations and excluding them without further justification undermines the credibility of the study. I believe that observations were excluded for good reasons. But those reasons should be explicitly laid out.”

Each observation which was excluded was first carefully examined. Missing data that was not reported to UZIS was, if possible, extracted from annual reports of individual hospitals. The data for missing values of other outputs and determinants was retrieved from alternative sources. If, in the end, DRG-adjusted patients or costs were missing, the observations was excluded.

Note in Figure 1.1 that extreme observations were not considered outliers and were left in the sample. However, if there was an obvious mistake stemming from the recording process, i.e. an observation did not fit into the time series of the hospital, it was excluded, leaving an unbalanced panel.

- “In the concluding remarks of Chapter 1, the author presents several recommendations for policy-makers. One of them states that ’the structure of management in non-profit hospitals reduces their efficiency.’ What would be a specific, more tangible recommendation on how to improve the structure of management in non-profit hospitals?”

A discussion was added to the conclusion of chapter 1 and to chapter 4
The results of the analysis suggest that nonprofit hospitals tend to be less efficient than their for-profit counterparts due to a different management structure. Most probably, the principal-agent problem is part of the explanation. The management of corporatized hospitals is contracted and is answerable to the shareholders and the supervisory board who may exert more pressure over them than public bodies over the management of budgetary organizations. For-profit hospitals are also subject to different accounting principles than their non-profit counterparts. The fact that budgetary organizations are not allowed to make profit and the earned amount must be put back, regardless of whether efficiently or not, contributes to lower scores of budgetary organizations in the analysis. However, Newhouse (1970) suggests that nonprofit hospitals have a different objective function which justifies their lower efficiency. For example, they may be situated in remote areas and thus their purpose is to guarantee access regardless of efficiency. Some wards in these hospitals are thus preserved even though they do not report a sufficient number of cases - in regions with alternative options these wards, and even hospitals, would be closed down.
This is not to say that all budgetary organizations should immediately be corporatized, but it is advisable that some elements from the more efficient for-profit model be transferred to nonprofit hospitals.

- “Finally, I was puzzled by the statement that ’Increasing utilization of magnetic resonance imaging and computed tomography units both in the Czech Republic and other Visegrad countries will increase the efficiency of the Czech healthcare system.’ In the Stylized facts section in the Appendix. On the one hand, advanced imaging may provide tremendous value to patients and providers due to its non-invasive nature. On the other hand, these diagnostic modalities are associated with potential harms (such as exposure to increased radiation in the case of computed tomography), as well as with considerably high costs. Moreover, the overuse (excessive use beyond what is necessary) of advanced imaging has been identified as a key driver of the rise of health care spending in many developed countries. Thus, I am confused

> about why the author believes that increasing the utilization of advanced imaging would improve the efficiency of a healthcare system, and what is the connection between the utilization of advanced imaging in other Visegrad countries and the efficiency of the Czech healthcare system."

Stylized facts in the appendix to chapter 4 should provide and overall picture on efficiency of the Czech healthcare sector. Efficiency increases due to increases in technology utilization is taken purely from economic and comparative perspective, i.e. that the average cost is the lowest upon maximum utilization. A discussion of adverse effect of overuse of technology in medicine was added. The last sentence was restructured. There is no connection between utilization in Visegrad countries and efficiency of the Czech healthcare system.

Changes to Chapter 2

Response to Olena Stavrunova

- "Set into the context of international literature on the price elasticity of the demand for healthcare."

Studies analysing the RAND Health Insurance experiment were already cited in the predefense version of the thesis, however, the text was edited so that it is obvious. Suggested papers dealing with price elasticity of the demand for healthcare were incorporated for the final version of the thesis.

- "Give institutional background about the implementation of reimbursement - are the areas affected by reimbursement on average wealthier, healthier, etc. than the rest of the country"? How did the reimbursement happen?"

Reimbursement was a political decision and thus could be applied only in hospitals controlled by regional (kraje) governments. Only regions governed by the Social democratic party (ČSSD) corporatized their hospitals. The electorate of this party reveals some characteristics, but given the fact that there are other hospitals as well in the area, it cannot be applied to the whole region. Reimbursement was based on a donation contract that a patient opted to sign in order to receive reimbursement or another form of agreement resulting in no need to pay copayment on spot for the patients (except for Plzeňský kraj in which the patients had to pay and were reimbursed within 2 months). Explanation of institutional background was added into the text.

- "The paper needs heterogeneity analysis. Is it possible to estimate the effect of reimbursement for different types of care (e.g. urgent vs. elective), or for different types of patients (young vs. old)? Elasticity of demand for health care varies by the type of care and there is heterogeneity across individuals as well."

Czech general hospitals analysed in this essay treat a wide spectrum of patients. The patients are of different age, sex, income, etc. Due to universal coverage and the third-party-payer system, the structure of patients treated is, however, homogeneous across hospitals both in treatment and control groups. A heterogeneity analysis of inpatient user-charges for a different structure of patients and types of care serves as motivation for further research since the available dataset does not allow for it.
The analysis for different age cohorts was carried out in Votápková & Žílová (2016a) which analyses outpatient user charges. A consistently insignificant effect was discovered across age cohorts. However, the user charge for ambulatory visits was even smaller than the user charge for an inpatient day. Price elasticity for different types of care as well as age cohorts was also carried out in Votapkova and Zilova (2016b) which analysed generic substitution in pharmacies finding out a different effect for drugs against chronic and acute illnesses. It was found that the elderly are price sensitive and prefer cheaper generics. When being prescribed a drug against acute illness, the patient does not opt for a cheaper generic but in case of drugs against chronic illnesses, the longer the generics is available, the more the probability increases that the patient chooses the generics, although at the beginning the patients rather chose the original. It suggests that transactions costs are too high for occasional drug users, but a regular drug user recognizes benefits of saving overtime when opting for generics. I believe that these results could be transferred also to the different types of inpatient care, however for a sufficiently large amount of co-payment. This discussion as well as motivation for further research was added to the last paragraph of conclusion.

Resposne to Andrea Menclová

- "The estimated effect in the thesis combines two mechanisms (1) User charges reduce the length of stay per patient and (2) User charges incentivize patients to move away from hospitals that effectively impose them to hospitals where charges are reimbursed. In the national scheme, mechanism 2 would not exist. Hence, chapter 2 potentially overestimates the effect of national user charges on utilization on inpatient care."

Indeed, if we admit that mechanism (2) exists, the real national effect would be even smaller, because under a comprehensive reform, we lose the substitution channel. But given the Czech context of the reform, we can assume that mechanism (2) is nearly non-existent as the costs of selecting a hospital further from home outweigh any savings from reimbursement of the small fee. For instance, a train ticket from Ceske Budejovice to Plzen, which are 140 km apart, costs 140 CZK. The substitution effect is thus not expected to play a role, even for longer hospitalizations. Besides, when a hospitalization is longer, a patient appreciates visits and personal support from family members whose travel costs are not imperceptible either. Traveling to a distant hospital is thus expected only for diversified treatment, which is however not a substitute to treatment in the nearest hospital. In addition, the competition variable is insignificant too. Thus, we assume that the estimates received from the natural experiment are indeed a good proxy for what similar charges would achieve if imposed at the national level. The discussion was added to sections 2.1. and 2.3.

- "However, there is also the possibility that patient selection across hospitals (in response to user charge reimbursement vs. not) is quantitatively important. I personal do not find that very likely given the small amount of these user charges but the author herself keeps revisiting this possibility, e.g. when writing: 'It was assumed that if the patients are price-elastic, they opted for user-charge reimbursement if they had a chance' (p.50) and 'one chooses a hospital in which one receives reimbursement for user charges.' (p.79) If that is the case, the small effects of user charges observed in chapter 2 are an overestimate of the effects of a national policy which estimates this strategic selection across hospitals."

The assumption about price elasticity was clearly defined. That is, given the small amount of user charge and the transaction costs of traveling to a different hospital, we do not expect the patients to be price elastic to this amount of user charge. Such an effect would likely play a role for a higher level of user charges.

- "I was a little confused by the author' s discussion of a substitution effect (first introduced on p. 42) and whether this is her attempt to get at the selection issue described above. In particular, she writes that she included 'a dummy variable acknowledging that the hospital is situated in a region where there is at least one hospital where patients could get reimbursement to account for a substitution effect.' (p.42). However, if people had better outside options and hence were more price elastic, I would expect the effect of user charges to vary depending on the presence/proximity of reimbursing hospitals. Hence, an interaction term between the proxy for user charges and the presence of nearby reimbursing hospitals would seem in order. Given that user charges are already proxied by an interaction term in a D-in-D model (post*treat), it would be easier to estimate the model on subsamples stratified by the presence of nearby reimbursing hospitals."

A robustness check on a subsample of hospitals stratified by the presence of nearby reimbursing hospitals was carried out (included in the appendix). The results of the robustness check confirm the results of the main analysis. The interaction term was insignificant in all model specifications. The effect of the reform is thus even weaker for a stratified sample than for the main analysis in which the interaction term was significant at least for some specifications. Should a substitution effect take place, the interaction term would be more significant here than in the main analysis. This robustness check thus confirms the absence of outside options under the analysed set-up. Yes, the attempt was to get to the discussion about the two potential effects, showing that no. 2 is non-existent. The text was re-structured so that it is clear.

- "However, that would also beg the question whether regional boundaries are a good measure of proximity; i.e. whether other hospitals in the same region are necessarily the most relevant outside options for any given patient."

The best outside option depends on the exact address of the patient which we cannot retrieve from the dataset. Of course, the patient may travel to a hospital in a different region. However, in some regions the possibility of reimbursement was restricted only to the patients living in that region, thus if there was a substitution effect, it should be the strongest for reimbursing hospitals within regional boundaries. The problem of institutional boundaries was explained in the institutional set-up of the reform.
Consistent with the explanation for the comment above, we do not expect user charge reimbursement to be the reason for travelling to a different hospital. The dummy variable "competition" was included to show that there is no substitution at play at all. I acknowledge that the definition of regional boundaries is a minor limitation though.

- "Overall, my vote would be for a relatively simple approach which acknowledges the possibility of mechanisms no. 1 and no. 2 as described above and hence suggests that the estimated effects of user charges are an upper bound of the effects a comparable national policy would be expected to have. I would then also add that mechanism no. 2 was, however, likely quite limited in practice as any cost savings from reimbursement could easily be outweighed by costs from selecting a hospital further from home."

I agree. But I also added the discussion of the competition variable which supports the non-existence of mechanism no. 2.

- "On a related note, it would be nice to emphasize just how low these charges were. ... I would recommend also comparing the amount to the prices of some basic items people routinely buy. From an international perspective, it is remarkable how much resistance a user charge can face when a night in a hospital under the care of highly-qualified medical specialists' costs individuals the same amount as, e.g. a moderately-priced meal at a restaurant or a pack of cigarettes."

Done.

- "The discussion of results in Chapter 2 often confuses economic vs. statistical significance. For example, when the author repeatedly writes that she finds 'marginal' effects, it is sometimes not clear whether this means 'small', 'marginally statistically significant, or both. For the most part, I would focus primarily on economic significance (i.e. the size of the interaction coefficient) and, to avoid confusion, I would call the estimated effect 'small' rather than 'marginal'. While statistical significance varies across model specifications, this is not very surprising with the relatively small sample size, and the point estimates on post*treat actually stay within a tight range (2.7-4.5 %) across models. "

Your interpretation was correct, sometimes it was economic, sometimes statistical significance and sometimes both, the same for marginal. It will be clarified. 'significance' will be used only for statistical significance, while economic significance, i.e. the size of the interaction coefficient will be denoted as 'small'. The discussion about the tight range will be set within the context of the sample size.

- "To put the author's work in the context of prominent health economics studies on related topics, it would be good to add a reference to the RAND Health Insurance Experiment studying the effects of patient cost-sharing on the utilization of medical care (and inpatient hospital care specifically)"

The reference was added together with other references suggested by other reviewers.

- "Minor: the D-in-D interaction coefficient is the key focus of the chapter. I would therefore put it in the first row of Table 2.2 to give it more prominence."

Done.

Response to Michal Horný

- "Given that the imposed user charges were billed to patients per day of stay, the dependent variable does have face validity. However, there is also a considerable incentive misalignment in this situation due to the principal-agent relationship between patients and providers. Although user charges were meant to reduce patients' moral hazard, their ability to influence the length of (inpatient) stay is limited. The potential of user charges to influence one's demand for the number of inpatient days lies mainly at the frontline decision whether to seek (inpatient) care at all or not. However, it is pertinent to keep in mind that most patients lack the knowledge to make an informed decision whether to get admitted and therefore rely on their provider's recommendations. Additionally, a non-negligible number of patients are admitted to a hospital unconscious or at least at a state of mind with limited ability of rational decision-making. Once admitted, the length of stay is driven primarily by each patient's health condition. Patients may request to be discharged earlier than recommended by their care-providing team, but such instances are relatively rare. These reasons are likely important factors of the small, marginally-significant effect found in the analysis. Moreover, providers acting as patients' (imperfect) agents have an incentive to extend the length of stay to increase own revenues. While this incentive is likely relatively weak in the Czech healthcare system (providers working at hospitals in Chechia are usually salaried), hospital management may encourage physicians to err on the 'safer' side and keep patients admitted longer than necessary. Moreover, this incentive may have become stronger with the abolition of user charges if hospitals management and providers knew that keeping patients admitted longer would have no negative financial impact on patients. Consequently, the dependent variable may be endogenous (as explicitly stated in Chapter 1: 'inpatient days [. . .] may bias the results due to possible endogeneity born by the length of stay . . . .' Thus, I encourage the author to consider the number of admissions as an alternative dependent variable in the analysis instead of the number of inpatient days."

I admit that the patient has a limited ability to influence one's length of stay in the hospital if the clinical conditions require it. This is of course what we do not want to reduce. The presence of user-charges should reduce the excess length of stay, i.e. additional examinations are carried out in one day except for postponing it to the following day. In this respect, the patients may insist on the providers to be more efficient and reduce excess length of stay. It is only partially true that providers/managers would have an incentive to artificially increase the length of stay. Even though still reimbursed by historical prices, a hospital must fulfill certain performance criteria in order to receive a percentage of the reference year's reimbursement. Specifically, if a hospital fulfils 97 % of inpatient production of the reference year measured through the DRG, it receives approx. 111.2 % of the budget of the reference year. If the production exceeds 100 %, a hospital is not reimbursed for this extra care.
Endogeneity was a problem for the analysis in chapter 1 but is not a problem for Chapter 2. In Chapter 2, we are interested in the total inpatient care provided and artificial access of the length of stay, be it induced by a provider or a patient, is exactly what we want to reduce through user charges. A proposed robustness check using admissions as the dependent variable would answer a different question than what the essay aims to and thus will be left for further research. In the current analysis, we are interested in the total amount of inpatient care provided, thus the total number of inpatient days perfectly proxies the subject of interest.

- "I applaud the author for including a section on the Validity of the difference-in-differences (DiD) methodology. DiD is a strong quasi-experimental research design that allows for causal inference under certain condition. However, whether all assumptions were satisfied ought to be tested. A critical assumption is that trends in outcome performance before intervention are parallel between treatment and comparison groups. The author inspected this assumption visually using a chart of aggregated data into two time-series. Ryan *et al.* (2015) provide a helpful checklist of the most important DiD methodology assumptions and the way they should be assessed. The parallel trends assumption ideally should be tested analytically using data of individual hospitals (not aggregated into time-series). I encourage the author to review the guidance on DiD by Ryan *et al.* (2015) and perform the recommended tests."

Ryan *et al.* (2015) was reviewed, however, based on discussions in the literature, a parametric test of the parallel trend assumption was carried out through the Leads and Lags model.

- "In section 2.2 Data, the author describes an assumption that 'children and the elderly have a higher price elasticity of demand for health care services and thus are more likely to ask for reimbursement if they have a choice.' I agree with the plausibility of this assumption for the elderly, but I am skeptical about the assumed price-elasticity of children's demand for health care given that the user charges for inpatient stays of children were likely paid for by their parents, and their potential influence on the number of inpatient days was likely minimal. How would relaxing this assumption affect the presented analysis?"

This was a pure reference to the literature. The statement was rephrased so that it is clear. The presented analysis would not be influenced anyhow. The productive population of Zlin region represents additional observations to the control group, that is 51 % of Zlin hospitals. The remaining 49 %, i.e. care provided to the children and elderly, were excluded, regardless of whether the children and the elderly actually asked for reimbursement, i.e. if they are price elastic or not. The data for the Zlin hospital was restricted in order not to lose these observations for the analysis since only a part of it is feasible to be used as a control group.
Based on the comments of other reviewers, the assumption of price-elasticity of specific age groups was thoroughly discussed.

- "The study was set up as an intent-to-treat analysis, which I believe was the right approach. Identifying whether patients were aware of the possibility of user charge reimbursement and if so, whether they acted upon it, would prove challenging in a retrospective analysis. Additionally, some patients in the pre-intervention period could have been exempted from the requirement to pay user charges if they had satisfied certain conditions (e.g. income below a certain threshold). How would the result change if the study estimated the average treatment effect (ATT) on the treated instead?"

There is no doubt that the patients were aware of the possibility to receive reimbursement since it was such a vivid political issue and a centrally planned policy step. Moreover, the hospitals were instructed to offer the possibility of reimbursement to the patients. Thus, even if we assume that lower educated (thus most probably low income) individuals do not follow politics, they were offered the possibility in the hospital. The individuals who qualified for the user-charge reimbursement based on the income threshold prior to the reform are a minority. I do not think that excluding these individuals from the sample would influence the results much.

Response to Martin Gregor

- "The chapter focus on the intensive margin in the patient's choices (quantity) but not on the extensive margin (self-selection into hospitals). Hence, to frame results, I would welcome a more detailed discussion of patients' self-selection into hospitals."

See response no. 1 to Andrea Menclová. Given the low level of user charges, we do not expect that that regional hospitals where patients could receive reimbursement represent an outside option for patients. The transaction costs to travel to another hospital exceed the low level of user charges. The observed effect is thus pure effect on quantity (intensive margin) that at the same time represents the true national effect, i.e. if the reform was applied on all hospitals in the country. The discussion was added into the text.

- "The analysis is carried out within a standard difference-in-difference setup. Among others, the DiD estimator requires that differences between treatment and control groups are time-invariant. Can we rule-out that the payment policy changes in the hospitals is not correlated with other policy/organization changes that are not included in the covariates? What about a falsification test which would replace the outcome variable by another variable that would not be affected by the policy change but could be potentially affected by policy/organization changes?"

Given the socio-economic situation and political decisions at the time, we can assume that there was no confounding effect on the number of patient days other than the reform of user-charge reimbursement that would cause an omitted variable bias in the DiD model. Thus, the effect of the reform was most probably plausibly exogenous. Unfortunately, the available dataset does not allow for a falsification test which would replace the outcome variable by another variable that was affected by the policy change.

- "For visual inspection of the parallel trends assumptions, what about reporting annual changes instead of levels (see Figure 2.2.)? Finally, one might also consider extending robustness tests, e.g. to try to control by systematic variation by using multiple control groups."

In figure 2.2. levels were changed to annual changes. As an additional robustness test, the dataset was stratified by the presence of a nearby reimbursing hospital. However, data availability limits the number of additional robustness test.

Changes to Chapter 3

Resposne to Olena Stavrunova

- “Include literature looking at the composition of financial portfolios near and after retirements, on the rates of wealth drawdown and on the effects of health and financial shocks on these rates (papers suggested in the report)”

Additional references suggested in the report as well as other references were added.

- “I missed some convincing argument why BMI and Life Satisfaction would be good instruments for health in the savings equation. First, I would expect body mass index and savings to be jointly determined by some common factors such as patience and self-control. I would expect that the relationship between health and life satisfaction runs in a different direction than in the equation in chapter 3, i.e. it is more plausible that good health increases life satisfaction, not the other way round, in which case, life satisfaction is also an endogenous variable”

It is always challenging to find appropriate instruments and then also to decide on variables which sufficiently satisfy exclusionary restrictions requirement for the structural equations. In order to decide the final specification of the structural equations, correlation conditions were checked. BMI and life satisfaction are not correlated at all; thus, a spurious correlation, i.e. a common factor influencing both variables, is not expected either. The causal relationship from life satisfaction to health was motivated by Ricketts *et al.* (2013) who assume job satisfaction to be positively related to health. We reject the endogenous relationship from health to life satisfaction because being healthy does not necessarily mean being happy. Healthy people may sometimes not realize the true value of life, i.e. the fact that they should be happy because they are healthy, thus they worry about other things. I can imagine that the endogenous relationship, i.e. healthy status would cause the people to be satisfied and happy, for the people who have overcome serious health issues. On average, however, I do not assume endogeneity to apply. This explanation was added to the text.

- “I am surprised that age is not included in the set of covariates in the Savings equation. We are looking at individuals between 50 and 60 years - when health starts to deteriorate very quickly, but also major financial decisions are made, e.g. increasing saving rate, increasing share of liquid assets in the household financial portfolio anticipation of retirement. It is essential to very flexible control for age in this equation to separate the effect of health shocks from other factors driving financial decisions in anticipation of retirement. I am also quite confident that the unexpected sign on subjective life- expectancy in this equation is because age is omitted from the model.”

Age was originally included in the Savings equation too but was insignificant and thus was excluded for the final model. Note, that the sample covers only the pre-retirement segment, i.e. 50-60 years. Thus, for the savings equation, it is probably very homogenous, i.e. everyone wants to form savings shortly prior to retirement. The age difference of 10 year prior to retirement is too narrow to be deterministic. This explanation was added to the text.

- "It would be good to see a longer discussion on how the effect of health on savings should be interpreted, after the simultaneity issue is solved by the instrumental variables. Is it the effect of negative health shocks that emerged in the pre-retirement period or rather is it the effect of a permanent ill-health that impeded accumulation of human capital and lead to low income and wealth over the life-cycle?"

From the dataset we cannot retrieve whether the individual had felt sick long before the interview or whether his/her health status is caused by a recent shock. Coile & Milligan (2006) show that health shocks cause changes in the financial portfolio and decumulation of wealth obviously because the sick want to pay to get well. A permanent ill-health is thus expected to disable the individuals to form savings in the first place. The inability to distinguish between these two may bias the results by the fact that if the health shock happened immediately prior to the interview, the respondents may not have responded to the shock with decreased savings yet. The bias is however not expected to be large but must be kept in mind while interpreting the results of the analysis. Thus, should the bias take place, the observed effect is underestimated. That is, without the bias, the positive effect of health on current financial assets would be even stronger. A short note was added to the text both to the data section and the results.

- "Finally, I think the chapter should explain why a cross-sectional data was used to study this essentially dynamic process. Most recent studies of wealth dynamics in retirement and pre-retirement rely on panel data which can separate individual fixed effects from the effects of changes in health in wealth. If this data set is the only source of data suitable to study the relationship between health and savings in the Czech Republic this should be explained in the paper."

The SHARE dataset Wave 5 interviewed individuals born in 1962 or earlier. Individuals interviewed in earlier waves are not the same. Our restriction for a homogenous pre-retirement segment, i.e. individuals 50-60 years of age very much restricts the possibility of a panel data model. A discussion was added to the text.

Response to Andrea Menclová

- "I am a bit concerned about the 'perceived life expectancy' variable for a few reasons. First, it is not clear to me whether it is self-reported or calculated from aggregate population statistics (e.g. by gender and age). Second, I am surprised by its low value in Table 3.1. Are only 56 % of 50-60-year-olds expected to be alive 10 years later? The minimum value of zero is also somewhat concerning (but I understand that the sample may include a terminally-ill individual and if this is based on self-reports. . .) Third, the author finds a negative effect of this variable on savings. While certainly plausible, I think it would be good to recognize that it is surprising (i.e. contrary to the author's expected sign)."

Perceived life expectancy is a self-reported variable. It expresses the probability with which the respondent expects to live in 10 years from the date of the interview. Thus, the mean in Table 3.1. expresses that on average a respondent expects to be alive 10 years later with 56 % self-perceived probability, not that 56 % of 50-60-year-olds expected to be alive in 10 years from the date of the interview. That is, each respondent had to report one's own expectations whether sh/e is alive in 10 years. The variable will be discussed in a more detail in the revised version of the dissertation. The sign contrary to the expectation will be discussed too. Most probably, the individual with a higher perceived probability expects to stay also in the labor market for some more time and thus expects to accumulate additional wealth. Note that the sample covered 50-60 year-olds, thus still a very young cohort to be active in the labor market without any serious difficulties.

- Table 3.1. Some of the values may deserve more thought/checking. For example, why is the youngest individual 26 when the sample should include 50/60 year-olds?"

I checked the re-checked the dataset as well as the Share Wave 5 documentation. It states that 'The target population for the baseline samples consists of all persons born 1962 or earlier having their regular domicile in the respective country.' Thus, when restricting the dataset to the pre-retirement segment of population, only the upper bound of persons older than 60 was used. It was a mistake not to check errors in the dataset beyond the lower bound. As many as 95 observations in the sample were born after 1962 thus being younger than 50 years of age. Although most of the people are not younger than 40 and thus would not distort the dataset too much, 9 observations report being younger than 40, including 5 being under 35. All observations below 50 years of age were excluded, and the analysis was re-run for the final defense.

- "Is the minimum value of education really 1 year?"

As many as 47 observations from the sample did not finish 8 years of education. The value of 1 appear in 5 cases only, 9 observations report 3 years of education or less. These seem to be outliers stemming from a wrongly specified question to the respondents (e.g. the respondent understood the question as additional years beyond basic school, etc.). All observations under 3 years of education were excluded.

- "What does the mean value of 1.43 for gender mean; i.e. are there more men or women in the sample?"

Yes, there are more women than men in the sample. Value 1 = female, 2= male. Sorry for a missing explanation in the text. It was improved for the final defense. Also, I recoded the variable into a dummy female = 0, male = 1.

- "Would it make sense to use BMI categories (underweight, normal weight, obese), rather than a continuous measure since the relationship with health status is not necessarily monotonic?"

For the final defense, the variable was recoded. Three categories were created, such that

Normal + underweight	bmi $< 25 = 1$	362 observations
(only 6 observations had bmi < 18.5, i.e. underweight		
Overweight	$25 \leq$ bmi $< 30 = 2$	565 observations
Obese	bmi $> 30 = 3$	366 observations

- "Finally, I am wondering if it would be worthwhile to recode the long-term illness variable into absence (1/0)."

For the final defense, the variable was re-coded to have the same interpretation as the variable 'self-perceived health status', i.e. absence of long-term illness = 1, otherwise 0.

- "Tables 3.2 and 3.3: As a general rule, tables of results should be self-explanatory (without reference to the main text). Tables 3.2 and 3.3 are currently difficult to understand without referring to the main text and this could be fixed if they are more carefully labeled. For example, use more informative titles than Model 1 and Model 2 vs. Model 3 and Model 4 and specify, what the dependent variable is prominently at the top of each panel (rather than in small font at the bottom)."

Done.

Response to Michal Horný

- "The author treats health status and savings as endogenous variables while assuming that other variables such as education, income, assets, perceived life-expectancy, partner's education, age, BMI and life-satisfaction are exogenous. With the only exception of age, which is clearly exogenous, exogeneity of the other variables is not obvious. A more detailed discussion of why each variable in the model can be treated as exogenous would make for a stronger argument about the conclusion of the study."

Finding an appropriate exogeneous instrument is always a challenge. If the exogenous variables are indeed exogenous, they are not correlated with random error values in the second stage regression. Sargan test and LM overidentification test show that all exogenous instruments are exogenous and thus uncorrelated with the model residuals. Specifically, the Sargan test here accepts the null hypothesis that the model is not overidentified. To carry out the LM test, we regress residuals extracted from the structural savings equation on all instruments. Since the null hypothesis is accepted, the model is not overidentified."

- "Self-perceived health status was recorded on a 6-point (0-5) Likert scale. As such, it is an ordinal variable but was treated as continuous in the analysis. Please elaborate on how this approach affects the interpretation of the results."

Estimating the first stage using probit or logit is not necessary because in 2SLS the consistency of the estimates in the second stage are not dependent upon specifying the correct functional form in the first stage (Kelejian, 1971). The more so that we are interested in the interpretation of the structural equation for savings.
Likert scale may be used in the analysis both as a continuous variable and a discrete choice of answer options depending on the purpose. In the analysis we consider numeric values since we are interested in how scores differ with respect to the independent variables. It is generally acknowledged that likert scale be used as continuous in parametric regressions if it has at least 5 point (the more, the better as Wu & Leung (2017) show). It allows us to believe that the interval between points are approximately equal. However, if we were interested in the response pattern across subgroups, then likert scale would be treated as a discrete variable and a log-linear or multinomial logit model would be more appropriate. A discussion will be added into the text.

- "In section 3.2 Data, the author mentioned that 'imputed variables were preferred due to many missing observations in the original dataset.' Please elaborate on which variables were imputed and how were the imputed values calculated."

Except for gender, all variables used in the analysis were imputed. Imputations are not my own but were done by SHARE themselves. Details how the imputation procedure was carried out is provided in Börsch-Supan (2016). To put it short, the variables with a low prevalence of missing values were imputed by univariate methods, such as hot-deck and regression imputations. Monetary variables with a high number of missing values were imputed using a fully conditional specification (FCS) method which preserves the original correlation structure of the variables(van Buuren *et al.*, 2006).
Elaboration was added to the text.

Response to Martin Gregor

- "I recommend being precise about the use of instruments in 2SLS (p.56): 'This approach ensures that error terms in the dependent variable are uncorrelated with the independent variables.' The lack of correlation is ensured only for a perfect instrument but the exclusion condition that is required for a perfect instrument obviously cannot be tested. I would also recommend fixing the interpretation of the overidentification test (p.63): 'Sargan test and the LM overidentification test confirm that all exogenous instruments are exogenous and thus uncorrelated with the model residuals.' In fact, the test assumes that at least one instrument is valid and tests for the validity of all other instruments. Hence it is not a panacea."

Sentence on p. 56 was changed and a footnote added such that:
'This approach ensures that error terms in the dependent variables are uncorrelated with the independent variables due to instrumenting.[9]'
The paragraph on p. 63 was fixed such that:
'Sargan test and the LM overidentification test show that the exogenous instruments are valid. Specifically, the Sargan test here accepts the null hypothesis that the model is not overidentified. The LM test assumes that at least one instrument is valid and tests for the validity of all other instruments. To carry out the LM test, we regress residuals extracted from the structural savings equation on all instruments. Since the null hypothesis is not rejected, the model is not overidentified.[10]'

Changes to chapter 4

Response to Andrea Menclová

- Some expressions should perhaps be used more cautiously or avoided. For example, the author often claims (throughout the thesis and again in Chapter 4) that her empirical analysis 'proved' a certain point. Econometric models never 'prove' a result - at most, they 'fail to reject' it. Hence, I would replace 'prove' with softer statements such as 'suggest', 'endorse', 'corroborate', etc."

Terms such as 'prove' and 'confirm' were replaced by softer statements.

- Other terms are sometimes used in Chapter 4 casually, without being properly defined. These could be misleading. Examples include: 'equal care', 'optimal level', and 'fair financing'. Normative statements on fairness, equity or equality should be treated with great caution."

[9] A perfect instrument satisfies the lack of correlation and the exclusion condition, the latter of which however cannot be tested. Finding a good instrument is thus always a challenge.

[10] An R-squared insignificantly different from 0 and an F test already reveals the like, however, they assume normality. Thus, the LM test is more appropriate.

Definitions were added and clarified.

- "As a minor comment: The author discusses corruption on p. 77. While corruption could lead to the phenomenon observed, there are also other explanations for hospitals' over-utilization of expensive capital - see for example Newhouse (1970) on the objective function of a non-profit hospital or various studies on ex-post moral hazard."

The text dealing with corruption in capital purchases was edited. However, objective function of non-profit hospitals (Newhouse, 1970) is definitely not responsible for overpriced purchases of capital in the Czech environment. Rather it is the third-party payer problem under which the management is not motivated to negotiate good prices or at least not buy visibly overpriced units. Moreover, the management is often rewarded for buying this overpriced equipment by the seller. Some of these cases have even been brought to court in the past.

Response to Martin Gregor

- "This chapter deserves minor editing and in some places a bit more extensive discussion of alternative channels, mechanisms, and policy options. For example, when discussing DRG, I would appreciate a clarification of the following statement (p.72): 'An alternative to the well-functioning DRG mechanism would thus look at local demographics, size, mission of the institution but at significantly higher costs than the DRG system.' What are the criteria for a well-functioning DRG? How exactly would the alternative work?"

The criteria for a well-functioning DRG is that the same production in different hospitals costs the same. No alternative to case-mix adjustment can guarantee this condition as effectively as the DRG mechanism. However, comparison of results of Votápková & Šťastná (2013) and Mastromarco *et al.* (2019) suggests that some of the adjustment is absorbed by local demographics etc., even though not perfectly.