THE EFFICIENCY OF ACUTE CARE HOSPITALS IN CANADA

THE EFFICIENCY OF ACUTE CARE HOSPITALS IN CANADA

By LI WANG, M.A, M.SC

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2019) Hamilton, Ontario

(Health Research Methodology)

TITLE: The Efficiency of Acute Care Hospitals in Canada

AUTHOR: Li Wang, M.A. (Nanjing University, Nanjing, China), M.Sc. (McMaster University, Ontario, Canada)

SUPERVISOR: Dr. Michel Grignon

NUMBER OF PAGES: 131

LAY ABSTRACT

A hospital is technically efficient if it uses its resources (its budget) to get the most in terms of quantity (number of stays) and quality of care it can. A hospital can be more or less technically efficient for reasons independent of its control (typically, because of the environment in which the hospital operates) and efficiency is a value-neutral measure. This thesis aims to: 1) estimate the technical efficiency of acute inpatient care in Canada and identify the potential factors that influence the level of efficiency achieved by a given hospital; 2) uncover the existence of possible geographic clusters of efficiency (hospitals that are close geographically are also close in the efficiency scale, something called spatial spillovers in the literature) in Canada.; and 3) examine the role that size plays in the variation of technical efficiency among small and rural hospitals across Canada.

The major findings are: 1) hospital output could be increased by 24 percent with the same resources by eliminating technical inefficiency; 2) There is a substantial and significantly positive spatial spillover effect on the efficiency of acute inpatient care: being close to an efficient hospital increases the efficiency score of a hospital, everything else being the same; and 3) The level of technical efficiency of small and rural hospitals across Canada is low overall and, perhaps surprisingly, larger rural hospitals are among the least efficient: among small hospitals, scale does not yield economies of resources.

ABSTRACT

Improving hospital efficiency is a critical concern for health care managers and policy makers. Hospital technical efficiency is measured as the ratio of what quantity and quality of care is produced to what could be produced given the level of resources available to the hospital (its budget). What a hospital should produce given the resources at its disposal is called the "production frontier". In order to improve hospital performance, health policy makers need knowledge and information about how well the hospitals they fund are utilizing the resources they receive. I apply Data Envelopment Analysis, a non-parametric technique, to administrative data on hospitals in Canada to produce the "technical frontier" and get insight into the variation of technical efficiency of acute hospitals at the Pan-Canadian level (except for the province of Quebec, which does not report its data on hospitals in a way that would make them comparable to the rest of Canada). DEA is preferred to the alternative method of stochastic frontier for the following reasons: DEA does not require to impose a specification on the production function of hospitals (for which theory is clearly lacking), and it allows the analyst to estimate a multi-output frontier (a stochastic frontier would have to weight arbitrarily the value of quantity versus that of quality of care in hospitals, whereas the DEA approach generates these weights from the data). Efficiency scores are serially de-correlated using a bootstrap technique and then entered as the dependent variable in regressions to identify the main factors of efficiency or inefficiency.

Specifically, this thesis aims to: 1) estimate the level of technical efficiency of acute inpatient care in 35 teaching hospitals, 54 large hospitals and 90 medium-size hospitals respectively in Canada and identify the potential factors that have influence on technical efficiency; 2) uncover and measure the existence of possible spatial spillovers of hospital efficiency in Canada and examine its potential determinants while taking into account the interaction between hospitals by means of spatial regression; and 3) examine the technical and scale efficiency of the 229 small and rural hospitals across Canada (outside Quebec), as well as estimate the impact of institutional and contextual variables on hospital technical and scale efficiency respectively.

The major findings are: 1) hospital output (combination of number and quality of stays; quality being measured as the inverse of in-hospital mortality) in Canada could be increased by 24 percent with the same resources by eliminating inefficiency. Highly efficient teaching hospitals benefit from producing care under favourable environments. Higher efficiency could be achieved by increasing cooperation within the health system and making more post- acute care beds available to both large and medium hospitals; 2) There is a substantial and significantly positive spatial spillover effect on the efficiency of acute inpatient care (elasticity of 0.3): Canadian hospitals are clearly complements to each other and work in networks much more than in competition. The hospital size (the number of beds), the percent of transfers between acute hospitals, and the percent of patient transfers to home care are the main drivers of efficiency among acute hospitals in Canada while controlling of the dependence between hospitals; and 3) Among small hospitals, the average output orientation technical efficiency on all types of services is 54% at the current input-output mix. To improve their technical efficiency, small hospitals should provide with more home care facilities to discharge their patients to (so-called Alternative Level of Care patients) and strengthen their cooperation with larger, urban hospitals. Small hospitals are scale inefficient, specifically, rural hospitals could reduce their size by 34% on average (around 6 acute beds) to achieve the optimal size. The study also found that the spending on diagnosis tests and the nursing as the percentages of total hospital spending (cost shares) are positively and significantly related to the scale efficiency.

ACKOWLEDGEMENTS

I would like to express my deepest gratitude to Dr. Michel Grignon for his comprehensive guidance, numerous discussions and constant support during the course of my training. I am grateful for his profound knowledge and insight that have brought me to the field of efficiency evaluation in the health sector and to where I am today. This work would not have been possible without his tireless dedication.

I am deeply indebted to my supervisory committee members, Dr. Jerry Hurley and Dr. Hsien Seow, who have been so supportive during this tough journey of PhD period. However, I cannot forget every moment you were supporting me by giving beneficial comments which really improved this thesis a lot.

I also want to extend my special thanks to Canadian Institution of Health Information (CIHI) for providing me the rich data and answering me many questions related to the data. I am extremely grateful for the CIHI researcher team and its members including Sheril Perry, Xi-Kuan Chen, Alison Ytsma, Sara Alin and Katerina Gapanenko for their invaluable comments and excellent editorial help, especially on my chapter II of the thesis.

Also, I would like to deliver sincere thanks to my friends for their initial encouragement and continued support.

Finally, special thanks go to my husband, Kai, for his encouragement, love, and support over the years, and also to my lovely son, Jerry, for his understanding and encouragement. I am also grateful for the blessing and support from my parents and my brother. I dedicate this thesis to them.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

FIGURES

CHAPTER I

CHAPTER IV

TABLES

CHAPTER I

CHAPTER II

CHPATER III

CHAPTER IV

LIST OF ABBREVIATIONS AND SYMBOLS

DMU - Decision Making Units
DRG - Diagnosis Related Groups
OECD - Organisation for Economic Co-operation and Development
HHI - Herfindahl–Hirschman Index
RIW - Resource Intensity Weight
CIHI – Canadian Institution for Health Information
CMG – Case Mix Group
MRI- Magnetic Resonance Imaging
PCI - Percutaneous Cardiac Intervention
HSMR - Hospital Standardized Mortality Ratio
DEA - Data Envelopment Analysis
SFA - Stochastic Frontier Analysis
CRS - Constant Returns to Scale
VRS - Variable Returns to Scale
TE – Technical Efficiency
PTE – Pure Technical Efficiency
SE – Scale efficiency
FEAR - Frontier Efficiency Analysis with R
LTC – Long Term Care
CMDB – Canadian MIS (Management Information System) Data Base
OMHRS – Ontario Mental Health Reporting System
MCC - Major Clinical Categories
MRDx - Most Responsible Diagnosis
LOS – Length of Stay
MIS - Canadian Management Information System
PCA - Principal Component Analysis
KMO - Kaiser–Meyer–Olkin
QALY - Quality Adjusted Life-Years
SRM - Spatial Regression Model
CAR - Conditional Auto Regressive
AHA - American Hospital Association
FY – Fiscal Year
GIS - Geographic Information System
OLS - Ordinary Least Square
SLM – Spatial Lag Model
SEM - Spatial Error Model
LM - Lagrange Multiplier
2SLS - Two-Stage Least Squares
IV - Instrumental Variables
GMM – Generalized Estimating Equations
EMS - Emergency Medical Services
MOHLTC - Ministry of Health and Long Term Care
CCHS - Canadian Community of Health Survey

CCR - Charnes, Cooper and Rhodes
BCC - Banker, Charnes and Cooper
NACRS - National Ambulatory Care Reporting System
ED – Emergency Department

DECLARATION OF ACADEMIC ACHIEVEMENT

I was the main contributor and the first author for all studies included in this thesis.

CHAPTER 1: INTRODUCTION TO HOSPITAL EFFICIENCY

1. Motivation of the Thesis

The pursuit of efficiency is a key policy objective of most health systems, along with equity and financial protection. Hospitals represent a significant proportion of health expenditure. In Canada, about 28.3% of total health expenditure is through hospitals (CIHI, 2017). A health system's efficiency is thus influenced to a great extent by the efficiency of its hospitals. A first motivation for the study of hospital efficiency is its impact on the efficiency of the health system as a whole. A study conducted in Australia has demonstrated that a 4% increase in hospital efficiency would result in a 1.9% increase in the overall efficiency of Australia's health care system (Commonwealth of Australia, 2009). Hospital efficiency has been measured at the national level in many countries (Mujasi et al., 2016; Magherini et al., 2016; Fragkiadakis et al., 2016; Jacobs 2001; Mutter et al., 2008; Kirgia and Asbu, 2013), or at the sub-national level in Canada (Gruca and Nath 2001; Bilodeau et al., 2004), but no study examines hospital efficiency at the Pan-Canadian level. Measuring efficiency of hospitals in Canada as a whole will therefore fill a gap in knowledge. However, it is not enough to measure the level of efficiency, one wants to know how efficiency can be improved: a first step in that direction is to study the determinants of efficiency, using the efficiency scores estimated in a first stage into a second stage econometric analysis. Such two-stages analyses are not frequent and none has been performed on Canadian data. Yet another motivation is methodological: the underlying model of efficiency analysis assumes that hospitals are decision making units in competition with each other, or, at the very least, with no interactions with each other: they all do the same thing, but separately, and more, or less, efficiently. However, it is clear that hospitals do collaborate with each other, and transfer patients from each other depending on their level of specialization. I therefore apply an original method to account for spatial spillover in the measurement of efficiency, as a way to enrich the measurement of efficiency with the recognition that hospitals are not standard decision-making units. The last motivation is part method, part policy: most studies of hospitals tend to exclude hospitals below a certain size, under the assumption that if a hospital is too small, it does not provide inpatient care. However, Canada, due to the existence of areas that are not densely populated, has a lot of small and rural hospitals, and it is important to understand the determinants of efficiency among these smaller hospitals, especially size: is there a scale effect among smaller and rural hospitals and could we provide recommendations regarding the optimal size of a rural hospital in Canada?

The thesis includes three papers. The first paper (chapter II) evaluates the technical efficiency of inpatient acute care in Canada and explores the impact of the environment in which the hospital works on the efficiency of acute inpatient care provided by that hospital. Canada seems to have gone further than most other comparable health systems in hospital specialization and concentration on acute care. In 2016, the last year for which we can compare hospital spending in Canada to that of other OECD countries, hospital spending represented approximately 26.4% of total health care expenditures in Canada, versus an average of just below 33.3% for OECD countries for the same year (only Ireland, Latvia and Mexico spent a smaller portion of their health care spending on hospital than Canada and 27 countries spent more, the largest share being 44.1% in Estonia) (OECD Health Statistics 2018, https://stats.oecd.org).

It is well-known that hospitals do not have full control of their production process and, as a result, of their technical efficiency: a large number of important decisions (admit, discharge, transfer, treatment course etc.) affecting efficiency are taken by physicians on the basis of the complexity and severity of each patient (Xenos et al., 2017), even though physicians are not paid or directly managed by the hospital's administration. Moreover, contextual effects such as the characteristics of the population (education, health-related behaviours) in the local area where the hospital is situated are important determinants of the efficiency of acute inpatient care delivered by the hospital (Matranaga et al., 2014; Mujasi et al., 2016; Fragkiadakis et al., 2016). Hospital managers have control over the (non-physician) cost structure, such as the average spending on administration as a per cent of total hospital spending or the quantity of nursing hours per weighted cases (Sultan et al., 2018, Kalhor et al., 2016), or the service mix of the hospital, such as the newborn cases as a proportion of total discharges (case-mix weighted) (Mujasi et al., 2016, Chowdhury and Zelenyuk, 2016). By benchmarking the performance of a hospital (efficiency of inpatient acute care) and examining its determinants, the generated conclusion will provide valuable insights to hospital managers who make operational decisions and to policymakers who may influence the external operating environment by regulations, reimbursement or by other policy measure.

The second paper (chapter III) examines the spatial interdependence of efficiency of acute inpatient care across hospitals in Canada. During the analysis of the first paper, we found that hospitals pose an interesting challenge to analysis of efficiency because the standard model of efficiency analysis posits that decision making units (DMUs) are substitutes who work in isolation, possibly with some level of competition. However, hospitals certainly compete but also complement each other and work in networks of care. Their final objective is not their bottom line but timely access to quality care for the community they serve and they will work in collaboration with nearby hospitals in order to do so. One specific case is collaboration between hospitals at various levels of specialization (secondary and tertiary) or other health care institutions such as home care and long-term care facilities. It was observed in the first analysis (efficiency and its determinant) that hospitals with better cooperation with other health care institutions tended to be more efficient. This motivated the exploration of the spillover effect of efficiency of inpatient acute care: does the efficiency of acute inpatient care in neighbouring hospitals affect the efficiency of the same type of care in a given hospital? A study conducted on German hospitals by Herwartz and Strumann (2010) has discovered a negative spatial spillover among German hospitals performance (a more efficient neighbour decreases efficiency), stemming from the fact that, under a prospective payment model based on diagnosis related groups (DRG) in Germany hospitals have become more competitors and less collaborators. More precisely, hospitals compete through selection and try to attract as many patients and as many low complexity patients as possible (Ridic et al., 2012, Norton et al., 2002). If competition were on cost only, it would have a neutral spillover effect. Because payment systems are less prospective in Canada (mixture of global budgets and DRG-based payment per episode, with varying power across provinces), we expect to see less incentives for hospitals to provide more services than necessary and more collaboration among hospitals. The ministry of health in each province is responsible for controlling medical expenditure through fixed global budget. The reimbursement exclusively takes place between the public insurer (the government) and hospitals and the patients do not participate in the reimbursement process. Therefore, the spatial interdependence in Canadian hospitals might be different from that in the Germany due to the different hospital care system. However, there is not any Canadian study focusing on directly testing how a hospital responds to the change of efficiency of its neighbor's hospitals. Thus, my study contributes to the literature by incorporating the spatial effect in the measurement of hospital efficiency.

I follow the approach suggested by Herwartz and Strumann (2010) to measure spillover effects on technical efficiency focusing on the acute inpatient care across hospitals. Similarly, Mobley (2003) and Mobley et al. (2009) examined whether hospital pricing are strategic substitutes or complements (do prices of neighbouring hospitals influence prices charged by a given hospital?) by adopting a spatial econometric model using data of 336 Californian hospitals for 1993 and 1999. They found a hospital responds to an increase in rival's prices by reducing its own price. Gravelle et al. (2014) applied the same approach to the question of quality spillovers in England and found that a rival's improvement in quality of care affects a hospital's quality in 6 out of 16 quality indicators, in which the rivals for each hospital are defined as the all other hospitals within a catchment area of 30-minutes travel time.

The approach builds on the fundamental concepts of spatial dependence and spatial autocorrelation as articulated in the "First law of geography" proposed by Waldo Tobler (1970, pp. 234-240): "everything is related to everything else, but near things are more related than distant things". The key tool of the approach is a spatial weighting matrix, that links each hospital in the data set to other hospitals in the data set, and provides, in each cell, either the distance between the two hospitals, or a binary variable indicating whether the two hospitals are located in the same geographic area (city, administrative region etc.). Distance thus measured between hospitals is hypothesized to determine the extent to which decisions by a hospital is influenced by decisions made by other hospitals. An econometric model tests the strength (and direction) of that influence, using the cells of the matrix as independent variables in a regression where "decisions" (price, quality, or "outcomes" in an efficiency study) are the dependent variables. The results are varied by using different types of spatial weighting matrix (Shahid et al., 2009). In this study, I tested the contiguity spatial matrix at the various distance thresholds (within a catchment area of 60km, 120km, 180km and 240km) to examine the impact of the catchment area on the magnitude of the spatial interdependence of hospital efficiency. It is standard in this method to set a threshold beyond which correlation is set at zero whatever the actual weight (if two hospitals are too distant, their correlation is assumed to be negligible). Additionally, I also tested other definitions of proximity, e.g., two hospitals can be geographically close but politically distant if they sit across a jurisdictional border (different provinces in Canada) and a hospital might influence a more distant hospital located in the same province than a closer hospital across the border. Hence, I tested using the binary spatial matrix (thresholds cut-off or whether within a province).

In this second paper, beside examining the spillover effect on hospital efficiency, I also examine whether controlling for the spillover effect changed the conclusions from the first paper on the determinants of efficiency: this is an important methodological contribution, as it can tell us whether neglecting spillover effects in estimations of hospital efficiency can lead to spurious findings on the determinants of efficiency.

Small and/or remote hospitals are excluded from the above two studies because they seem to serve a different purpose in the provincial health care system than large, medium and teaching hospitals. Small hospitals often act essentially as community health centres that support the triage of acute patients as well as post-acute care recovery of patients discharged from larger hospitals (it can also be said that local hospitals are protected by local politicians as they are often the main employer in rural areas). Some provinces (e.g. British Columbia) over-compensate small hospitals for the number of cases they treat because small hospitals cannot be expected to perform at the same level of efficiency or complexity as larger or urban hospitals. Countries with large numbers of small hospitals, such as Australia and the United States, also acknowledge the fact that small and remote hospitals cannot operate as large hospitals and fund small or remote hospitals through their global budgets or cost-plus reimbursement policies (Sutherland et al., 2013). Therefore, my third paper (chapter IV) is to examine the technical and scale efficiency among small and rural hospitals that typically serve sparsely populated residents in rural areas. I want to examine two key research questions. The first question is similar to that in my first two papers, that is to measure the average technical efficiency and its determinants. The second question is to examine whether scale matters among small hospitals: how small can a small hospital be, or should small hospitals be really small to fulfill their mission? Similar to the first two papers, the third paper also aims to explain the measures of performance by regressing the efficiency scores (technical and scale efficiency) against a variety of factors that might affect hospital performances.

It is well known that small hospitals suffer from a lack of economies of scale. However, it is important to measure to what extent this affects their performance. My findings, somewhat surprisingly, show that removing scale inefficiency would require downsizing some small and rural hospitals as well as upsizing some of them. By identifying the factors that are significantly associated with scale efficiency, beyond the reduction(increase) in size, small hospitals could find some alternative ways to make themselves more scale efficient.

In general, the three papers in my thesis provide a clear framework for policy implications to increase the overall efficiency of the small hospitals in Canada.

Leveraging the rich data that the Canadian Institute for Health Information (CIHI) collects systematically and harmonizes across provinces, this empirical work contributes to the existing literature in a number of ways: 1) This is the first ever study of hospital efficiency (acute inpatient care) at the Pan-Canadian level except for the province of Quebec; 2) I take advantage of the non-parametric DEA method to measure the multi-dimensional output of hospitals, quality of care being as important as quantity (case-mixed adjusted number of discharges); 3) I contribute to the application of an original method of measuring spatial inter-dependence to the question of efficiency spillovers; and 4) I apply the method of scale efficiency to the seldom studied set of small and rural hospitals, a crucial policy question in Canada.

2. Conceptual model of hospital efficiency

My thesis focuses on the technical efficiency (the chapter II and III and IV) and scale efficiency (chapter IV) of hospitals in Canada. In this section, I will illustrate some methodological issues common to all three chapters, including the scope of the study, data availability and the DEA model in the analysis. Especially, I would like to express my gratitude to Sheril Perry, Xi-Kuan Chen, Alison Ytsma, and Katerina Gapanenko from the CIHI who provided insight and expertise and drafted some parts of this section. Any errors are my own and should not tarnish their reputations.

Economists do not all agree on definitions of efficiency and we provide here the definitions we follow in our study, to avoid ambiguity and misunderstandings. Efficiency can refer to allocative efficiency (producing the combination of outputs that will maximize social welfare) or to productive efficiency (producing as much outputs as possible given the resources available, Farrell 1957; Leibenstein 1966). Making the assumption that hospitals are missioned by the single payer to produce a volume of treatments within the budget they are allocated, Canadian hospitals can be inefficient in two different ways: they can waste some of the inputs, something we call “technical inefficiency”, or they can use all units of inputs but in a combination that is not the lowest cost of all possible combinations of inputs to get the same level of outputs, something we call “cost-effectiveness inefficiency”. Ideally, we would want to separate technical from cost-effectiveness inefficiency in our analysis, but we did not have access to data separating unit price (cost) from quantity of resources and must be content with using global budgets as our measure of inputs. Under these circumstances, the budget of the hospital and our efficiency analysis is the supply- side efficiency analysis, as distinct from allocative efficiency, which obviously is on the demand-side.

In general, to measure the efficiency of something, one needs to determine three main components: decision-making unit (DMU), output and input.

2.1 Decision-making Unit

The definition of a Decision Making Unit (DMU) in the case of estimating hospital efficiency is not straightforward. It is important that all DMUs used in the model apply the same types of inputs to produce the same types of outputs; and that DMUs are so defined that it is in their power to make decisions about their inputs and outputs. It is also important, that DMUs are comparable to each other: they are peers.

Hospitals vary in size (from extra-large to very small), care provided (from all type of care to very specialized care), and location (from urban centers to rural and remote). Hospitals are also influenced by government and social factors that lie outside of the hospital’s control such as availability of discharge to home care and LTC, and overall population health characteristics. Moreover, there is no consistent definition of a *hospital* across Canada. Appendix 1 presents the definition of *hospital* in some provinces. Beside the category of small and remote hospitals, described in the introduction of the chapter IV, I distinguish three categories within the population of non-small hospitals:

Teaching hospitals

Teaching hospitals treat complex patients and have higher costs per patient as a result, which makes it difficult to compare them to other large non-teaching hospitals. Hence, the efficiency of the teaching hospitals should be estimated separately. Teaching facilities also require extra staff for the instruction of training medical staff.

Hospitals with concentrated services

Hospitals that provide more care in a specific area might be more or less efficient than other hospitals. This difference may come not from *how* hospitals deliver care but rather from *what type of care* they deliver. Concentration of services might be a result of political regional decisions and/or previous capital investments or types of population who use the hospital. Regardless of the reason, hospitals with concentrated services need to be excluded from the analysis as the concentration of services of a hospital could profoundly influence its comparability to other hospitals. Examples of concentrated services include pediatric, mental health, women's health, and facilities specializing in procedures such as cataracts, hernias, or joint replacement.

I use the Herfindahl–Hirschman Index (HHI) to identify hospitals with concentrated care. HHI is a measure used traditionally in economics to estimate the size of firms in relation to the industry and as an indicator of the amount of competition among them. To apply this approach to identifying specialty hospitals, I use the distribution of stays across major clinical categories (MCC) within each hospital, and calculate the HHI as the sum of squared market shares (here, the discharges of each MCC as a percentage of total discharges). Specialized hospitals are those with an HHI greater than 0.25.

Multi-site hospitals

Some hospitals operate as a network or have multi-sites. Ontario in particular has a number of hospitals that have more than one physical locations (20 out of 124 hospitals). The different sites work closely together and for our analysis, a transfer between two sites within a hospital should not be considered a transfer. As a result, we treat all hospitals within such a network as a single DMU. This, obviously, raises the issue of less formalized networks of hospitals, studied in chapter III (spatial spillovers).

Last, it should be noted that hospitals produce a variety of types of treatments, from emergency care to inpatient to day surgery and outpatient care. In this thesis (except for chapter IV), we focus exclusively on inpatient care by far the largest portion of a hospital budget, and one for which we can arrive at a reasonably well standardized definition of outputs (see next sub-section).

2.2 Hospital outputs

I will measure outputs as a combination of quantity and quality of care provided by the hospital. Quantity is measured as the number of discharges from acute care beds, weighed by the resource intensity weight (RIW) developed by CIHI, where each patient receives an RIW by the CMG+ methodology that groups patients by diagnoses, and relative resource use and adjusts for the presence of co-morbidities and interventions (CIHI, 2017). Using RIW to weigh stays is the most practical approach but RIW is associated with the CMG + groups which imply treatment for a certain condition rather than production of a specific service. This is the limitation of the analysis and I discuss its weaknesses below.

Hospital's choice of less expensive procedures

Using RIW to measure quantity of care will not benefit hospitals that choose less expensive procedures to treat patients. If two procedures exist to treat the same condition, one using more resources than the other, each procedure receives its own RIW. By using RIW, we will not penalize a hospital that would consistently choose the high RIW over the low RIW even though the high RIW might yield lower quality of life overall.

Hospital's decision between inpatient and outpatient treatment of the same condition

Hospitals should be encouraged to substitute very expensive inpatient care with less expensive alternatives (for example, a shift from inpatient to outpatient treatments). A hospital that uses technology to shift patients from inpatient to outpatient care, even though another provider within the system provides the outpatient care, should be deemed efficient. However, in our approach, such a hospital might be seen as less efficient, since it will not provide as many RIWs as a hospital that substitutes less outpatient treatment for inpatient. Cataract surgery, dialysis, chemotherapy, radiotherapy, mastectomy or tonsillectomy are just a few examples of outpatient treatments that used to be inpatient activities. "Over the past several decades, the number of nights Canadians spent in acute-care hospitals on a per capita basis has declined, while post-acute and alternative services provided in the home and community have grown" (Health Canada, 2011, https://www.canada.ca/en/health-canada/services/health-care-system/reports-publications/health-care-system/canada.html). Appendix 3 – Figure 1 shows the distribution of total hospital expenses in fiscal year of 2012-2013 in the teaching, large, medium and small hospitals.

Using RIWs may bias efficiency measures when a hospital shifts its *inpatient treatment to outpatient services*. Ideally, cost-weighted outputs should combine the number of treatments of inpatient, outpatients and chronic care. However, inpatient treatments and outpatient treatments are assigned different RIWs even though they may both have the same effect on patient outcomes. An implicit assumption in the cost-weighted activity is that a treatment with lower unit costs has lower quality than a treatment with higher unit costs. This is the limit of any (virtual) prospective payment scheme that is not fully powered (power is lower than 100%) because it is partially unbundled (Newhouse, 2004).

Another approach, followed, e.g., by Statistics Canada in their productivity study (Gu and Morin, 2014) is to weigh hospital activities based on their value to patients: they assigned the same RIW to both inpatient and outpatient activities within the same case type (case mix group). However, it is not perfect either, as it assumes implicitly that hospitals are always free to decide on whether to treat any case inpatient or outpatient. If they are not, and can only treat milder cases outpatient, assigning the same RIW to outpatient and inpatient hospital visits will penalize hospitals with more severe cases. There does not seem to be any perfect solution and I opted for restricting my analysis to the efficiency of inpatient acute care rather than hospital care as a whole.

The case of day surgery adds specific difficulties since what is performed as day surgery in one jurisdiction will be provided in a clinic in another. This happens frequently with investigative procedures (endoscopy) and select procedures (dialysis, chemotherapy, cataract, vasectomy). I analyzed data on day surgery. I restricted the definition of day surgery to cases involving moderate to general degree of anesthetic, thus excluding the cases requiring deep anesthesia. The intent is to restrict cases for inclusion to those occurring in what the financial data identifies as the day surgery unit. The results show that almost 90% of all applicable interventions that could be happening in day surgery are occurring in the day surgery setting. Over 99% of day surgeries are happening in groups where more than 80% happen in the day surgery setting. The assumption going into the analysis appears to be largely true: That the vast majority of interventions appropriate for day surgery are happening in the day surgery setting. As a result, there should not be too much variation across hospitals in the choice of providing some interventions in the day surgery setting as opposed to the inpatient setting, and excluding day surgery from the analysis will not bias the results (i.e., will not unduly penalize hospitals which provide more day surgery). There are, however limitations. There is a very small portion of major surgeries happening in the day surgery setting, including joint replacement and brain interventions. There are also some procedures that appear to be happening in both the inpatient and day surgery setting, including a variety of spinal, cardiac and reproductive organ procedures. There might be variations across hospitals in the frequency with which these procedures are provided as day surgery as opposed to inpatient setting and, on these procedures, my results will be biased. These procedures represent less than 0.6 % of total hospital spending though, and this should not affect our overall conclusions (Appendix 3 – Figure 2).

Patient transfers between hospitals

In some instances, more than one hospital treats a patient in the course of an inpatient treatment: hospitals cooperate with each other by sending their patients to a more appropriate hospital or accepting patients sent from other hospitals. These transfers make the hospital sector more efficient: for instance, some low-volume procedures are concentrated in a small number of hospitals in the province, because offering them in all hospitals would jeopardize quality (if a hospital performs a given procedure less than a certain number of times in a given year, it is not very good at performing it) and increase costs (there are fixed costs associated with performing a procedure). As a result, patients needed that procedure will be transferred to the concentration hospitals (who are likely large or teaching hospitals). For similar reasons (concentrating specialists improves quality and reduces costs), large and teaching hospitals concentrate certain types of care (mental health, pediatric care, women's health, heart etc.). As a result, medium (and small) hospitals will transfer cases needing those types of care to these large and teaching hospitals that have the capacity to treat such cases. In such cases, the patient is often transferred back to the sending unit. Transfers also take place to regulate flows: if a large hospital is overwhelmed at one time with cases in a given specialty, it can transfer some of its cases to a nearby hospital with available capacity. This saves costs to the system in the long-run, since, otherwise, all hospitals would have to support reserve capacity in case that would remain unused most of the time. Last, a large, medium or teaching hospital can transfer a patient to a smaller, less resource-intensive hospital for surveillance after an intervention: typically, the patient would receive high-technology care in the admitting hospital before being transferred to a medium hospital closer to home. This, again, saves resources to the system as a whole and improves quality (the patient's well-being is improved by recovering closer to their family).

To evaluate the efficiency of a hospital, we therefore need to account for these transfers by adjusting the RIWs. Currently, Canadian hospitals assign "Atypical" RIWs to patients with long hospital stay, those who were transferred, died, or left against medical advice. However, in our analysis we use "typical" RIW which excludes the atypical cases because these adjustments provide resource for potentially inefficient care (see details in appendix 2). As a robustness check, to ensure this choice did not drive our results, we run a sensitivity analysis using the normal RIWs which include the atypical weights (see chapter II for detail).

The output of a hospital cannot only be measured as the number of patients processed, but must take the quality of care received while in the hospital into account. We want to monitor what hospitals are providing their patients, not only how many they admit and discharge. CIHI collects many indicators to measure quality of care (CIHI, 2016). The most common indicators include the mortality rate (premature, avoidable, treatable), hospital re-admission (all-cause), hospital standardized mortality ratio (HSMR), in-hospital mortality rate for specific diseases, in-hospital sepsis rate, adverse event rate and nursing sensitive adverse events (medical, surgical).

It is not practical to use many indicators in measuring efficiency; therefore, a decision needs to be made on what indicators should be included in calculating the scores and what indicators will be used to regress hospital efficiency scores against system factors. It is important to make sure that the quality indicators selected to be included in the economic model are available for every hospital included in the analysis. More detail about these decisions is provided in the method section (output) of chapter II.

2.3 Hospital inputs

As already mentioned, I do not have access to information on volumes (hours of nursing for instance) and will use levels of spending (unit cost multiplied by volume) instead. This implies that inputs are measured as flows (what is spent in a given period) and not stocks (what has been accumulated so far), which makes a difference mostly for capital (I will measure amortization of installed capital rather than value of the stock of capital accumulated by a hospital). Hospitals spend their budget across three main categories of expenditures: Capital amortization (the annual depreciation of the stock of buildings, machines and beds, including those purchased the current year)[1], labour (wages of nurses, health care aids, support staff, etc., paid in the current year) and other operating expenses (food services, laundry, etc., also paid in the current year). Total hospital expenditure is commonly used as an input measures.

CIHI data allow me to distribute expenditures across functions (objective of the expenditure), but I do not use that information and treat spending as a whole to avoid the need to allocate shared hospital expenses like meals and laundry. More detail on this decision is provided in the method section (input) of chapter II.

In Canada, *physician' fees* are not part of hospitals' budgets but clearly contribute to what hospitals produce and are part of the resources they use. A strong case can be made in favour of

---

1 If a hospital makes an investment, e.g., buys a new Magnetic Resonance Imaging (MRI), CIHI does not count the whole amount of the MRI machine in that particular year as capital spending, but takes the first year amortization of the machine only into account.

including physicians' fees to the set of inputs used by the hospital, if hospital managers control the number of doctors who receive privileges in their hospital, and if they can control the number of patients they treat in the hospital. However, because I don't have strong evidence that this is actually the case (and it can be argued that doctors use the hospital as their workshop, leaving no control whatsoever to the hospital managers), and also because there is no data in Canada on the number of doctors having privileges in a given hospital or the number of patients they see, the physician's fees were excluded from this study.

The scope of the study is presented in Appendix 4 and Figure 3 (Appendix 5) outlines the concept model of the hospital efficiency in Canada.

3. Methodology for the determination of the frontier (calculation of efficiency scores)

3.1 Data Envelopment Analysis

Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) are the two main techniques to measure efficiency. In both approaches, the frontier is derived from information on all decision-making units, implying that efficiency scores are relative, as opposed to an absolute measure that would be based on an expert-based frontier (e.g., an engineer would tell how much outputs can be produced with a given quantity of inputs). The difference between the two methods lies in how the information on comparators is used: the SFA is essentially a regression method, and the frontier is determined by the average relationship between inputs and outputs (information from all decision-making units will influence the position of the frontier). The DEA uses information on neighbours only, i.e., decision-making units with similar inputs and the frontier is built in a piece-wise manner using linear programming techniques. I used the DEA for the following reasons.

(1) SFA relies heavily on *a priori* assumptions on the specification of the production frontiers and the random error distribution. Within a Canadian context, there is no widely accepted research that can be used as *a priori* assumptions on the functional form of the production process in acute inpatient care.

(2) The strong assumptions on the functional form and the error distribution make SFA vulnerable to model misspecification (O'Neill et al., 2008). Also, there is no way to test whether the *a priori* assumptions are appropriately specified.

(3) SFA cannot deal with multiple outputs for the production efficiency: as a result, some arbitrary assumption needs to be made to weigh the outputs relative to each other.

Unlike SFA, DEA is a flexible approach and it doesn't require *a priori* assumptions on specification because the model determines the efficient acute hospitals by comparing the actual practicing results of all other hospitals. i.e., if one hospital can produce more than another with the same level of inputs, the latter is not efficient. More specifically, one hospital is inefficient to some extent if a linear combination of two other hospitals observed in the data that would use the same level of inputs can produce more (Jacobs et al., 2006).

There are different ways to conduct a DEA analysis and decisions have to be made: The first decision is the orientation: with input-oriented DEA, the linear programming model is configured so as to determine how much the input use of a DMU could contract if used efficiently in order to achieve the same output level. In contrast, with output-oriented DEA, the linear programming model is configured to determine a DMU's potential output given its inputs if it operated efficiently as DMUs along the best practice frontier. We employ an output-orientation approach to DEA, based on the assumption that ministries of health are interested in getting closer to achieving their objectives given a fixed budget. The second decision is about returns to scale (either constant, CRS, or variable VRS). The latter encompasses both increasing and decreasing returns to scale. CRS reflects the fact that output will change by the same proportion as inputs are changed (e.g. a doubling of all inputs will double output); VRS reflects the fact that production technology may exhibit increasing, constant and decreasing returns to scale (See Figure 4)

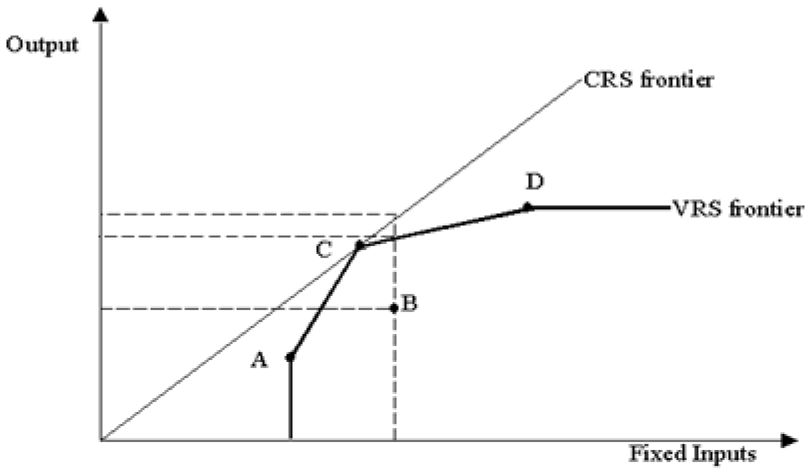

Figure 4: the DEA CRS frontier and VRS frontier in the simple one input-one output case

I follow the VRS approach, mostly because it is more flexible than the CRS: if returns are constant, the VRS will find it (e.g., DMU C in Figure 4), but if I impose a CRS assumption to data that are not CRS, the CRS will not find VRS (it will assume that DMUs A and D are inefficient in Figure 4). Also there is evidence of diminishing marginal returns in the health sector (adding more resources increases output by less) for instance the relationship between spending and health outcomes at the country level (so-called flat of the curve care, OECD, 2011).

Formally, the output-oriented DEA model under VRS specification for the $HOSPITAL_0$ is specified as the following mathematical formula:

$$min_{\emptyset,Z}\emptyset_0$$

$$Subject{:}\frac{-y_{r0}}{\emptyset_0}+\sum_{j=1}^{n}z_jy_{rj}\geq 0r=1,2,\ldots,m$$

$$x_{i0}-\sum_{j=1}^{n}z_jx_{ij}\geq 0i=1,2,\ldots,k$$

$$z_j\geq 0,j=1,2\ldots,n$$

$$\sum_{j=1}^{n}z_j=1;$$

where $y_{rj}$is the vector of outputs (r) for *HOSPITAL$_j$*, x_{ij} is the vector of inputs (i) for *HOSPITAL$_j$* and the values of outputs and inputs should be nonnegative. z_j represents a vector of weights attached to each *HOSPITAL$_j$* in the comparison group from the perspective of *HOSPITAL$_j$* (Charnes, 1978) and is determined by the above linear programming problem. $\emptyset$ is the efficiency score that measures technical efficiency (TE) of a hospital and satisfies $\emptyset\leq 1$. In this paper, we use Shephard's distance measure of output orientation technical efficiency, which is bounded between zero and one, i.e., a TE score of 1 indicates that the HOSPITAL lies on the production frontier representing relative technical efficiency. When TE is less than 1, the HOSPITAL is inside the frontier (e.g., DMU B in Figure 4), therefore technically inefficient, and the more inefficient, the lower the TE score.

The scale efficiency reflects the portion of inefficiency attributed to the scale of operations. If a hospital is scale efficient, any modification to its size (keeping the proportions of inputs constant) will result in a less efficient production. Scale efficiency is measured as the ratio of CRS technical efficiency and VRS technical efficiency scores (Cooper et al., 2007)

3.2 Limitations of DEA

The DEA method has some limitations that need to be addressed. First, DEA results are sensitive to high-performing outliers. If the outliers are undetected, the large random variation affecting a frontier DMU moves the entire frontier, which influences the estimates of all other DMUs. It thus has severe consequences for the analysis. This study employs the method proposed by Wilson to detect the outliers in DEA. The approach essentially adopts the influence function based on the geometric volume spanned by the sample observations and the sensitivity of this volume with respect to deleting suspicious observations from the sample (Wilson, 1993). The detailed mathematical description of the method can be found in Wilson (1993) or Fried et al. (2008). Sensitivity analysis therefore is commonly applied before the DEA to test how much influential observations affect the frontier and the efficiency of the other DMUs (see details in chapter 2, method section). The macro "ap" and "ap.plot" in the FEAR package are used for the detection of outliers.

Second, DEA measures efficiency relative to a deterministic frontier driven by the sample observations. However, the resulting efficiency scores incorporate random noise as well as inefficiency. This is because the observed input-output data would be normally subject to measurement error, omitted input or output variables, or one-off events: a DMU can be within the frontier on a given year due to a year-specific event (e.g., shortage of long-term care beds to discharge patients) rather than to structural inefficiency. These errors affect the measured efficiency scores. For hospitals on the frontier, this problem is acute since errors will not only affect efficiency measured for the hospital with the data error, but possibly other hospitals as well. To address this drawback, I use the smoothed bootstrap method developed by Simar and Wilson (1998) to correct the estimates of efficiency for random noise. This approach allows the calculation of so-called "bias-corrected" estimates of efficiency (hereafter referred to as robust estimates), and to calculate confidence intervals around those estimates. The fundamental idea in smoothed bootstrapping is that we assume that the probability distribution of estimated DEA efficiencies mimics the true but unknown population of efficiencies. Thus, if we draw with replacement a sample from the estimated DEA efficiencies, it will be like drawing a sample from the population itself. By sampling repeatedly from the observed DEA efficiencies we can construct an empirical sampling distribution for the DEA efficiencies of units. Then the empirical sampling distribution is used to estimate confidence intervals on the DEA efficiencies. The algorithm to Bootstrapping DEA using output orientation includes the following steps:

(1) Employ DEA to the original data to calculate efficiency scores (point estimate): $\widehat{\Phi}_j$

(2) Generate a random pseudo sample with replacement of size *n* from the empirical distribution of efficiency scores $\widehat{\Phi}_j$, providing $\Phi^*_{1b,} \cdots \Phi^*_{nb}$.

(3) Obtain a bootstrap set of pseudo-outputs$\left(y^*_{jb} = \frac{\widehat{\Phi}_j y_j}{\Phi^*_{jb}},\ j = 1, \cdots n\right)$using the ratio of the original efficient output level, i.e., the product of original efficiency score and the original input ($\widehat{\Phi}_j y_j$) and the pseudo-efficiency scores (Φ^*_{jb}).

(4) Adapt DEA to this new set of observations, composed of the pseudo-outputs (y^*_{jb}) from step 3 and the same set of inputs then calculate the bootstrapped efficiency scores ($\widehat{\Phi}^*_{j,b}$).

(5) Repeat steps 1-4 B times to generate a distribution of size B of bootstrapped scores for statistical inference.

The estimated bootstrap bias is calculated by B times Monte-Carlo simulations using

$$\widehat{bias}_j = \frac{1}{B}\sum_{b=1}^{B} \widehat{\emptyset}^*_{j,b} - \widehat{\emptyset}_j = \overline{\emptyset}^*_j - \widehat{\emptyset}_j$$

A bias corrected estimator of efficiency $\emptyset_j$ is $\widetilde{\emptyset}_j = \widehat{\emptyset}_j - \widehat{bias}_j = 2\widehat{\emptyset}_j - \overline{\emptyset}^*_j$

Smith (1997) notes that a well-specified DEA model will always overestimate efficiency; therefore, robust estimates are consistently lower than the point estimates. Because DEA is a descriptive approach that defines the efficiency frontier with actual units, some of them will have an efficiency score of 1, by construction. It is highly likely, though, that the true frontier is further out and that no unit actually achieves perfect efficiency, something that the smoothed bootstrap will generate. The macro "boot.sw98" in FEAR package is used for the smooth bootstrapping process.

Third, DEA is sensitive to the number of inputs and outputs with respect to the number of DMUs used in the analysis. Efficiency scores are likely to be overestimated if the number of DMUs is small relative to the number of inputs (Mohammad, 1998). Also, a large number of DEA inputs and outputs can result in an excessive number of DMUs lying on the frontier, which reduces the ability to identify inefficient DMUs (McCallion et al., 2000). Hollingsworth and Peacock (2008) recommend that the number of DMUs should be at least three times the total number of input and output variables, which is the case in all analyses presented in this thesis.

4. Analyzing the determinants of hospital efficiency (regressing the scores on factors)

4.1. The determinants of hospital efficiency

This section presents a summary of some sources of variation in technical efficiency between hospitals with respect to two aspects: one is what hospitals can manage and control and the other is the effect of the environment in which the hospital operates. The potential determinants of the hospitals were collected based on the literature. The underlying assumption of each potential determinant is s presented in Appendix 6 of chapter II.

4.2 Method of analyzing the determinants of hospital efficiency

A commonly employed method is the two-step Data Envelopment Analysis (DEA) approach. First, we estimate the hospital efficiency scores, using the non-parametric DEA approach that constructs the efficient frontier with the best performing observations of the sample. Second, these estimates are then regressed (e.g., Tobit or OLS) against a set of explanatory variables in an attempt to explain observed efficiency. However, Simar and Wilson (1998) note that existing two-stage studies do not account for the fact that DEA efficiency estimates are serially correlated and proposed the smoothed bootstrapping method (see above) to solve this problem. Since then, many researchers have adopted this methodology for use in hospital efficiency studies. Although bootstrapping in two-stage approaches has become a mainstream technique since 1998, some researchers pointed at its own problems. For example, Afonso and Aubyn (2007) noted that bootstrap estimates are based on a set of assumptions concerning the data generation process and the perturbation term distribution that may be disputed. Kirigia and Asbu (2013) and Jehu-Appiah et al. (2014) cited this methodological problem as a reason for avoiding bootstrapping in two-stage approaches. These limitations led Simar and Wilson (2007) to develop the double bootstrap procedure that enables consistent inference within DEA models estimating and explaining efficiency scores, while simultaneously producing standard errors and confidence intervals for these efficiency scores. In the study, I applied the procedures following Simar and Wilson's (2007) Algorithm 2. It consists of the following steps. Firstly, standard DEA efficiency point estimates are calculated (step (i) in Appendix 6). Secondly, truncated maximum likelihood estimation is used to regress the efficiency scores against a set of explanatory variables (ii). These estimates are then integrated into a bootstrap procedure that is similar to the smoothed bootstrap procedure of Simar and Wilson (2000) (iii). This bootstrap procedure allows to correct for bias (iv). Finally, the bias corrected scores produced by the preceding bootstrap are used in a parametric bootstrap on the truncated maximum likelihood (v-vi), thus producing standard errors for the regression parameters. Confidence intervals are then constructed, for the regression parameters as well as for the efficiency scores (vii). This procedure is described in more detail in Appendix 6 (drawn from Simar and Wilson, 2007, Algorithm 2).

5. Thesis outline

This thesis presents several empirical applications of econometric methods for the assessment of performance of hospital in Canada except for the province of Quebec. Specifically, the thesis is organised as follows: In chapter II, I estimate the average level of technical efficiency of acute inpatient care by using hospital activities together with measures of quality of care for the teaching, large and middle hospitals respectively. The determinants of efficiency of acute inpatient care are evaluated through a Tobit regression on scores estimated by the double bootstrapping method. Chapter III analyzes the same set of hospitals and further explore spatial spillovers of hospital performance (technical efficiency of acute inpatient care) by taking into account the interdependence among hospitals in Canada. The direction and the magnitude of the spillover effects, as well as the sources of the inefficiency are analyzed by the spatial lag model. Chapter IV focuses on assessing the performance of the small or remote hospitals in Canada. The technical and scale efficiencies of small hospitals in term of the acute inpatient care, as well as the overall services in Canada are examined. Chapter V summarizes the findings and provides concluding remarks.

References:

Afonso, A. and St. Aubyn, M., 2011. "Assessing health efficiency across countries with a two-step and bootstrap analysis." *Applied Economics Letters*, *18*(15), pp.1427-1430.

Gravelle, H, Santos, R, Siciliani, L 2014. "Does a hospital's quality depend on the quality of other hospitals? A spatial econometrics approach." *Regional Science and Urban Economics,* vol. 49, pp. 203–216. DOI: 10.1016/j.regsciurbeco.2014.09.005

Bilodeau, D., Crémieux, P.Y., Jaumard, B., Ouellette, P. and Vovor, T. 2004. "Measuring hospital performance in the presence of quasi-fixed inputs: an analysis of Quebec hospitals." *Journal of Productivity Analysis*, 21(2), pp.183-199.

Canadian Institute for Health Information (CIHI) 2016. *Health Indicators 2016*. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI) 2017. *National Health Expenditure Trends, 1975 to 2017*. Ottawa, ON: CIHI; 2017

Canadian Institute for Health Information (CIHI). 2017. *DAD Resource Intensity Weights and Expected length of Stay*. Ottawa, ON: CIHI.

Charnes, A., W. Cooper and E. Rhodes 1978. "Measuring the efficiency of decision making units." *European Journal of Operational Research*, 2: 429–444.

Chowdhury H, Zelenyuk V 2016. "Performance of hospital services in Ontario: DEA with truncated regression approach." *Omega*. 2016;63:111–22.

Commonwealth of Australia 2009. "Public and private hospitals: Productivity Commission research report. Canberra, ACT", Australia: Productivity Commission: Commonwealth of Australia, 2009.

Cooper W, Seiford L, Tone K.2007. *Data envelopment analysis a comprehensive text with models, applications, references and Dea-Solver software*. New York: Springer; 2007

Farrell, M. J. 1957. "The Measurement of Productive Efficiency." *Journal of the Royal Statistical Society,* Series A (General), CXX (1957), 253-81.

Fragkiadakis, G., Doumpos, M., Zopounidis, C. and Germain, C., 2016. "Operational and economic efficiency analysis of public hospitals in Greece". *Annals of Operations Research,* 247(2), pp.787-806.

Fried, H., Lovell, C. A. K., & Schmidt, S. 2008. *The measurement of productive efficiency and productivity growth*. London: Oxford University Press, 638 p

Gravelle, H, Santos, R, Siciliani, L 2014. “Does a hospital’s quality depend on the quality of other hospitals? A spatial econometrics approach.” *Regional Science and Urban Economics* vol. 49, pp. 203–216. DOI: 10.1016/j.regsciurbeco.2014.09.005

Gruca, T. S., and D. Nath. 2001. “The technical efficiency of hospitals under a single payer system: the case of Ontario community hospitals.” *Health care management science* 4(2): 91-101.

Gu, W., and S. Morin. 2014. “Experimental measures of output and productivity in the Canadian hospital sector, 2002 to 2010.” *In Measuring Economic Sustainability and Progress*, pp. 575-594. University of Chicago Press.

Health Canada, 2011, *Canada’s Health Care System*, https://www.canada.ca/en/health-canada/services/health-care-system/reports-publications/health-care-system/canada.html

Herwartz, H, & Strumann, C 2010. “Spatial Interdependence of Hospital Efficiency in Germany.” paper presented to Beitrage zur Jahrestagugng des Vereins Fur Socialpolitik 2010: Okonomie der Familie, Session: Public Policy Spillovers, No. F7-V2

Hollingsworth, B. and S.J. Peacock, 2008. *Efficiency Measurement in Health and Health Care*. New York: Routledge.

Jacobs, R., 2001. “Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis.” *Health care management science*, 4(2), pp.103-115.

Jacobs, R., P. C. Smith and A. Street. 2006. *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*. Cambridge, UK: Cambridge University Press.

Kalhor R, Amini S, Sokhanvar M, Lotfi F, Sharifi M, Kakemam E.2016. “Factors affecting the technical efficiency of general hospitals in Iran: data envelopment analysis.” *J Egypt Public Health Assoc*. 2016;91:20–5.

Kirigia, J.M. and Asbu, E.Z., 2013. “Technical and scale efficiency of public community hospitals in Eritrea: an exploratory study.” *Health economics review*, 3(1), p.6.

Leibenstein, H. 1966. “Allocative efficiency vs." X-efficiency".” *The American Economic Review*, *56*(3), 392-415.

Magherini, L., Lettieri, E. and Agasisti, T., 2016. *Evaluating technical efficiency and the effect of innovative healthcare delivery on Italian hospitals: A two stage DEA approach*. Documents from XXI Summer School “Francesco Turco.

McCallion,G.M., J.C.Glass, R.Jackson, C.A. Kerrand D.G. McKillop. 2000. “Investigating productivity change and hospital size: a nonparametric frontier approach.” *Applied Economics*, 32: 161–174.

Mobley, LR 2003. “Estimating hospital market pricing: an equilibrium approach using spatial econometrics.” *Regional Science and Urban Economics* vol. 33, no. 4, pp. 489-516. DOI: 10.1016/s0166-0462(02)00059-5

Mobley, LR, Frech, T, & Anselin, L 2008. “Spatial interaction, spatial multipliers, and hospital competition.” Economics Working Paper Series 07-08, Department of Economics, UC Santa Barbara.

Mohammad, R. Alirezaee 1998. “Sampling size and efficiency bias in data envelopment analysis.” *Journal of Applied Mathematics& Decision Sciences*, 2(1): 51–64.

Mujasi, P.N., Asbu, E.Z. and Puig-Junoy, J.. 2016. “How efficient are referral hospitals in Uganda? A data envelopment analysis and tobit regression approach.” *BMC health services research*, 16(1), p.230.

Mutter, R.L., Rosko, M.D. and Wong, H.S., 2008. “Measuring hospital inefficiency: the effects of controlling for quality and patient burden of illness.” *Health Services Research*, 43(6), pp.1992-2013.

Norton, E.C., Van Houtven, C.H., Lindrooth, R.C., Normand, S.L.T. and Dickey, B., 2002. “Does prospective payment reduce inpatient length of stay?”.*Health Economics*, 11(5), pp.377-387.

Newhouse, J.P., 2004. “Reimbursing for health care services.” *Économie publique/Public economics*, (13)

O’Neill L, Rauner M, Heidenberger K, Kraus M. 2008. “A cross-national comparison and taxonomy of DEA-based hospital efficiency studies.” *Socio-Economic Planning Sciences* 42: 158–189

Organisation for Economic Cooperation and Development (OECD), 2011, *Health at a Glance 2011: OECD Indicators*. http://www.oecd.org/els/health-systems/49105858.pdf [24 September 2013].

Ridic, G., Gleason, S., & Ridic, O. 2012. “Comparisons of health care systems in the United States, Germany and Canada.” *Materia socio-medica*, 24(2), 112.

Sultan, W. I., & Crispim, J. 2018. “Measuring the efficiency of Palestinian public hospitals during 2010–2015: an application of a two-stage DEA method.” *BMC health services research*, 18(1), 381.

Sutherland, J.M., Crump, R.T., Repin, N. and Hellsten, E., 2013. *Paying for hospital services: A hard look at the options*. (CD Howe Institute Commentary No. 378).

Shahid, R, Bertazzon, S, Knudtson, ML, & Ghali, WA. 2009. “Comparison of distance measures in spatial analytical modeling for health service planning.” *BMC Health Services Research*, vol. 9, no. 1, pp. 1-14, DOI: 10.1186/1472-6963-9-200

Simar, L. and P.W. Wilson, 1998. “Sensitivity analysis of efficiency scores: how to bootstrapping nonparametric frontier models.” *Management Science*, 44(1): 49–61.

Simar, L., & Wilson, P. W. 2000. “A general methodology for bootstrapping in non-parametric frontier models.” *Journal of applied statistics*, *27*(6), 779-802.

Simar, L. and P. W. Wilson 2007. “Estimation and inference in two-stage, semi-parametric models of production processes.” *Journal of Econometrics*, 136: 31-64

Tobler W., 1970. "A computer movie simulating urban growth in the Detroit region." *Economic Geography*, 46(Supplement): 234-240.

Wilson, P. W. 1993. “Detecting outliers in deterministic nonparametric frontier models with multiple outputs.” *Journal of Business and Economic Statistics*, 11(3): 319–323.

Xenos, P., Yfantopoulos, J., Nektarios, M., Polyzos, N., Tinios, P., & Constantopoulos, A. 2017. “Efficiency and productivity assessment of public hospitals in Greece during the crisis period 2009–2012.” *Cost Effectiveness and Resource Allocation*, 15(1), 6.

Appendix:

A1. Definition of a hospital

BC Hospital Act

"Hospital" means a nonprofit institution that has been designated as a hospital by the minister and is operated primarily for the reception and treatment of persons

(a) suffering from the acute phase of illness or disability,

(b) convalescing from or being rehabilitated after acute illness or injury, or

(c) requiring extended care at a higher level than that generally provided in a private hospital licensed under Part 2;

Alberta Hospital Act

"Hospital" means an institution operated for the care of diseased, injured, sick or mentally disordered people;

Saskatchewan Health Legislation (The Facility Designation Regulations)

3 For the purposes of section 10 of the Act:

(a) the following categories of facilities are established:

(i) addiction treatment centre;

(i.1) complex care centre;

(ii) health centre;

(iii) hospital;

(iv) residential treatment centre;

(v) special-care home; and

(b) with respect to the hospital category, the following subcategories are established:

(i) community or northern hospital;

(ii) district hospital;

(iii) regional hospital;

(iv) provincial hospital;

(v) rehabilitation hospital.

6 If a facility or part of a facility is designated as a **community or northern hospital**:

(a) it must provide to in-patients and out-patients:

(i) medical services;

(ii) basic radiography and laboratory services;

(iii) emergency stabilization services;

(iv) observation and assessment services; and

(v) convalescent care and palliative care; and

(b) it may provide any of the following:

(i) out-patient surgical services;

(ii) obstetrical services;

(iii) services mentioned in clauses 5(f) to (o).

7 If a facility or part of a facility is designated as a **district hospital**:

(a) it must provide to in-patients and out-patients:

(i) medical services;

(ii) basic radiography and laboratory services;

(iii) emergency stabilization services;

(iv) observation and assessment services;

(v) convalescent care and palliative care; and

(vi) obstetrical services; and

(b) it may provide any of the following:

(i) surgical services;

(ii) services mentioned in clauses 5(f) to (o).

8 If a facility or part of a facility is designated as a **regional hospital**:

(a) it must provide to in-patients and out-patients:

(i) medical services;

(ii) basic radiography and laboratory services;

(iii) fluoroscopy and computerized tomography diagnostic services;

(iv) emergency stabilization services;

(v) observation and assessment services;

(vi) convalescent care and palliative care;

(vii) surgical services;

(viii) obstetrical services;

(ix) intensive care services; and

(x) specialty medical services in the areas of internal medicine, general surgery, obstetrics and gynecology; and

(b) it may provide any of the following:

(i) specialty medical services in areas including, but not limited to, orthopedics, ophthalmology, urology and otolaryngology;

(ii) rehabilitation services;

(iii) services mentioned in clauses 5(f) to (o).

9 If a facility or part of a facility is designated as a **provincial hospital**:

(a) it must provide to in-patients and out-patients:

(i) medical services;

(ii) basic radiography and laboratory services;

(iii) fluoroscopy and computerized tomography diagnostic services;

(iv) emergency stabilization services;

(v) emergency and trauma services;

(vi) observation and assessment services;

(vii) convalescent care and palliative care;

(viii) surgical services; and

(ix) specialty medical services in the areas of internal medicine and general surgery; and

(b) it may provide any of the following:

(i) interventional radiology, magnetic resonance imaging, nuclear medicine and hemodynamic laboratory services;

(ii) intensive care services;

(iii) obstetrical and gynecological services;

(iv) intensive neonatal and pediatrics services;

(v) specialty and subspecialty medical services and surgical services;

(vi) rehabilitation services;

(vii) services mentioned in clauses 5(f) to (o).

10 If a facility or part of a facility is designated as a rehabilitation hospital, it must provide to in-patients and out-patients:

(a) rehabilitation services; and

(b) one or more of the services mentioned in clauses 5(f) to (o).

Manitoba Hospital Act

"Hospital", except in sections 57 to 60, means

(a) a hospital in Manitoba that is designated as a hospital by regulation under Subsection 113(1), and

(b) a hospital or facility outside Manitoba that

(i) is approved by the minister for the purpose of this Act, and

(ii) is licensed or approved as a hospital by the governmental licensing authority in the jurisdiction in which the hospital is located;

Ontario Hospital Act

"**Hospital**" means any institution, building or other premises or place that is established for the purposes of the treatment of patients and that is approved under this Act as a public hospital.

Nova Scotia

(f) "**hospital**" means a building, premise or place approved by the Minister and established and operated for the treatment of persons with sickness, disease or injury and the prevention of sickness or disease, and includes a facility, a maternity hospital, a nurses' -residence and all buildings, land and equipment used for the purposes of the hospital, or means, where the context requires, a body corporate established to own or operate a hospital, or a program approved by the Minister as a hospital pursuant to this Act or any other Act of the Legislature

A2. CIHI - Overview of CMG+ grouping and RIW methodology

CMG+ Grouping Methodology

The purpose of the CMG+ grouping methodology is to identify clinically relevant patients and recommend adjustments for clinical severity. This is achieved using the clinical information provided in the CMDB data, mostly using the most responsible diagnosis (MRDx) and intervention codes. Initially, the MRDx is used to assign each patient to one of the 21 Major Clinical Categories (MCCs) which are primarily dependent on the part of the body affected, or specialty.

If interventions related to the MCC assigned are reported, the patient is assigned to a CMG within the surgical partition; otherwise, they are assigned to a CMG within the medical partition. If the patient has more than one intervention that is related to the MCC, the most resource intensive one is used to assign the CMG. In the medical CMGs, the specific CMG assignment is again based primarily on the MRDx, and often further adjusted for additional services or indicators such as infection or complicating diagnosis. For example, within the medical partition, there are two CMGs for Arrhythmia to adjust for those cases with coronary angiogram, the presence of. In both surgical and medical CMGs, other diagnoses (comorbidity) and interventions (flagged interventions) are used to adjust cost estimates for the severity of the patient.

RIW methodology

The RIWs (resource intensity weight) is an estimate of relative resource use based on the assigned CMG+ groups and factors. The cost estimated are developed based on 56 primarily large or teaching hospitals in ON, AB, and BC with about 50% of the data coming from ON facilities, 40% coming from AB facilities, and the final 10% coming from BC. The expected resource use is assigned based on CMG+ and factors, but subsequent adjustments are made for atypical (death and transfer cases) and for exceptional long-stay cases. Transfers and deaths present unique challenges as the weight assigned is based on length of stay (LOS). The LOS for these cases tends to be variable as the patient does not receive full treatment, but these patients are often the more complicated cases requiring surgery or with complex medical conditions. Almost 5% of cases are considered as having exceptional length-of-stay measures, and LOS adjusted to the RIWs reflects the actual acute care total cost. For the purposes of evaluating hospital efficiency, it is not ideal to adjust cases for exceptional LOS as this is often an indication of inefficient care. Hospitals with a high portion of long stay care are more likely to be small rural hospitals with poor health system support such as additional mental health, rehabilitation, home and-or long term care services. To address over-resourcing of long stay cases, for hospital efficiency, no long-stay RIW adjustments will be included.

Table 1 provides RIW calculations for the current CMG+ methodology and the proposed adaptations for the Hospital Efficiency study. RIW calculation method is determined by chosen based on presence (or absence) of factors, whether the patient was a transfer patient, and whether the patient was a long stay* patient. These three issues are independent, therefore there are 8 (2*2*2) possible scenarios.

Table 1: Summary of RIW and Hospital Efficiency(HE) Adjusted RIW calculation by type of case

		Not transfer or death		Transfer or death	
		Not long stay	Long stay	Not long stay	Long stay
No Factors	CMG+ Methods	Base RIW	Base RIW + (Base Long Stay Per Diem*(LOS-ELOS))	Base Per Diem*LOS	Base RIW + (Base Long Stay Per Diem * (LOS-ELOS))
	HE Methods	**Base RIW**	**Base RIW**	**Base Per Diem*Regional Transfer LOS**	**Base Per Diem*Regional Transfer LOS**
Factors	CMG+ Methods	Base RIW *Factor Effects	(Base RIW * Factor Effects)+(Base	Base Per Diem * Factor Effects *LOS	Base RIW * Factor Effects +(Base Long Stay

			Long Stay Per Diem * Factor Effects *(LOS-ELOS))		Per Diem * Factor Effects * (LOS-ELOS))
	HE Methods	**Base RIW** * Factor Effects	**Base RIW *** Factor Effects	**Base Per Diem *** Factor Effects * **Regional Transfer LOS**	**Base Per Diem *** Factor Effects ***** **Regional Transfer LOS**

*Long stay means that LOS is greater than trim point. Trim is based on CMG, age group, and factors (if any).

A3: Model coverage

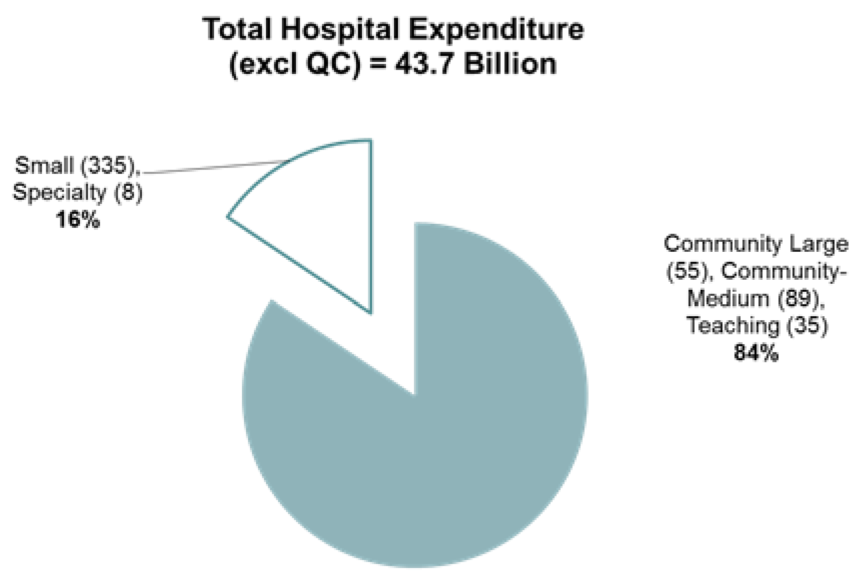

Figure 1: What portion of total expenditure is covered in the sample of hospitals?

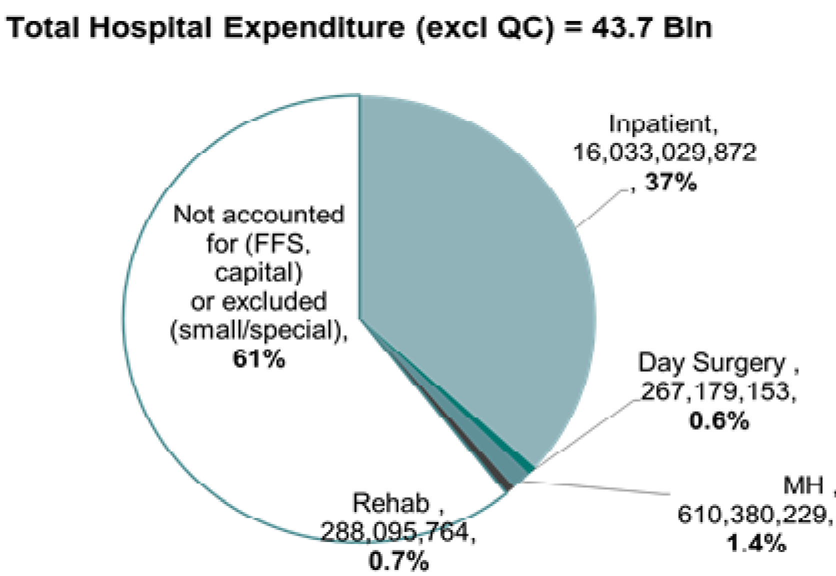

Figure 2: What proportion of total hospital expenditure is accounted in the project

A4: Study Inclusion Criteria

Coverage in all applicable data holdings (OI, HSP, DAD, NACRS, CMDB, OMHRS

OI – Organization Index – Organizational Structure and Relationships

HSP – Health System Performance – Outcome Measures

DAD – Discharge Abstract Database - Inpatient care and Day Surgery

NACRS – National Ambulatory Care Reporting System - Day Surgery from select provinces

CMDB – Canadian MIS (Management Information System) Data Base - Cooperate level financial data

OMHRS – Ontario Mental Health Reporting System - Ontario non-acute care mental health beds.

Exclusion Criteria and Implementation Plan

- ***Rehabilitation Cases and Hospitals***

Rehabilitation hospitalizations have been reported to different data holdings.

DAD acute care facilities data

DAD as or rehabilitation, the National Rehabilitation Reporting System (NRS)

The reporting structure varied by province/territory. A few hospitals from BC and AB have dual reporting to both NRS and DAD. Data quality team is investigating the magnitude of dual reporting. CMDB team cannot estimate the cost for rehabilitation hospitalizations reported to NRS. We would suggest not including rehabilitation in our hospital efficiency analysis.

- ***Quebec data***

Due to differences in the way the Quebec collects their data, it is not easy to obtain completely comparable input (hospital costs) and output measures (RIW) with the Quebec data. It is believed that this information in something that could be considered with specific adjustments considered for Quebec.

- ***Mental Health Hospitals***

Mental health patients in acute-beds will be included but we will not include mental health specialized hospitals and psychiatric hospitals. If RIW cannot be satisfactorily mapped to OMHRS data, the all mental health case will need to be excluded.

- ***Other Specialty Hospitals***

Specialty Hospitals were identified as outlined in

    - Identified using HHI index based on MCC distribution by hospital.
    - Flag facilities with primarily one specialty: Cataract, Hernia, Mental health and hoteling (MCC 20).

- Surgical, medical, oncology and obstetric hospitals will be flagged as well.

Study Limitations

- Exclusion of physician billing
- Exclusions
  - Rehabilitation
  - Quebec Data
- Only a regional assessment of social determinants of health
- Lack of provincially and regionally assigned expectations of services
- Assumptions associated with CMG and CMDB methodology

A5: Outline of components to be included in hospital efficiency study

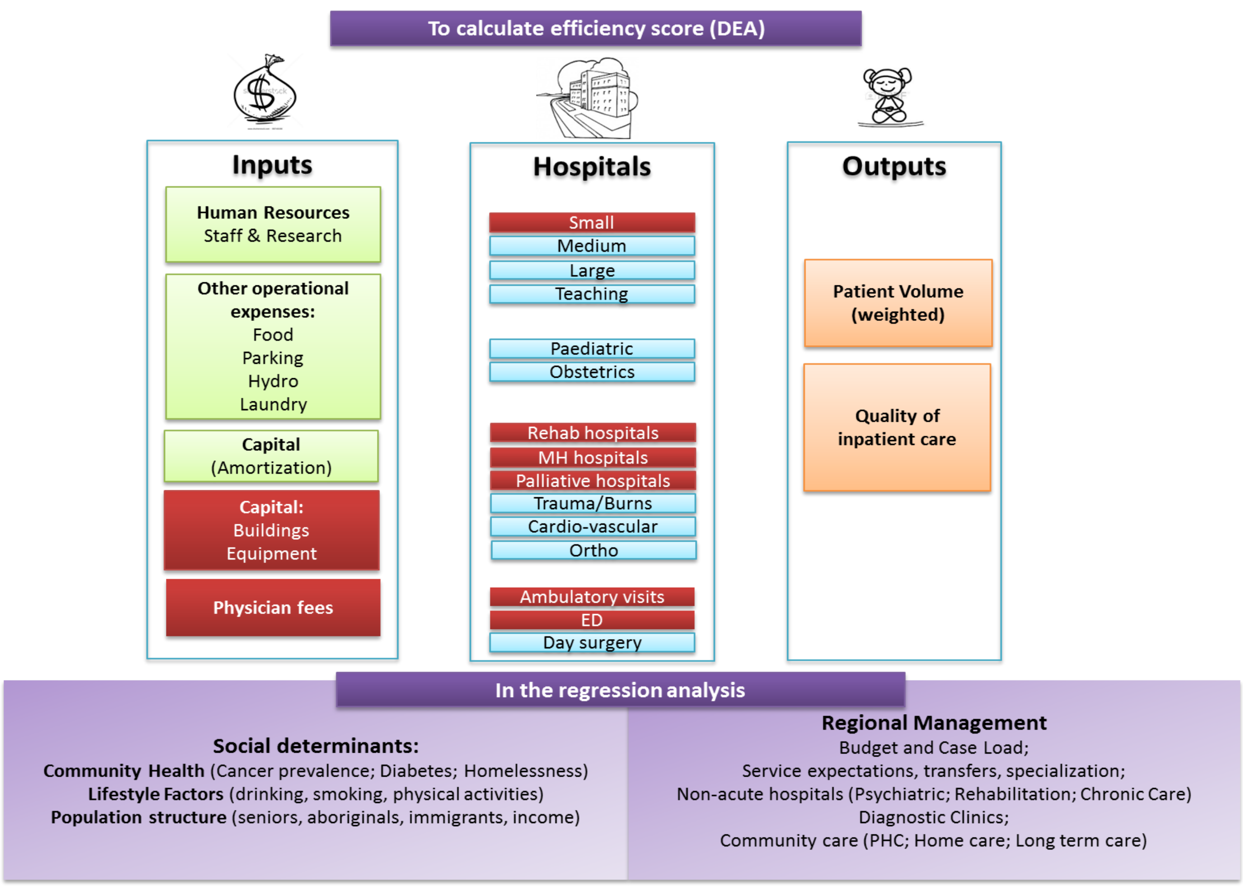

Figure 3: Outline of components to be included in hospital efficiency study

A6: Double bootstrap procedure algorithm 2

The seven steps of the double bootstrap algorithm are as follows:

i) A DEA output-orientated efficiency score $\hat{\delta}_i$ is calculated for each farm, i.e. the following program is solved for i=1,…,n (CRS case):

$$\max_{\lambda,\hat{\delta}_i} \quad \hat{\delta}_i$$

$$\text{subject to} \quad -\hat{\delta}_i y_i + Y\lambda \geq 0$$
$$x_i - X\lambda \geq 0$$
$$\lambda \geq 0$$

where y_i and x_i are respectively the original output and input matrices of the *i*-th farm; Y and X are respectively the original output and input matrices of the sample; λ is a nx1 vector of constants; the DEA score $\hat{\delta}_i$ is bounded by one on the left: $\hat{\delta}_i \geq 1$ for i=1, …n. For a specification under VRS the additional constraint $\overline{\mathbf{1}}\boldsymbol{\lambda} = \mathbf{0}$ is added, where $\overline{\mathbf{1}}$ is a vector of ones.

ii) Maximum likelihood is used in the truncated regression of $\hat{\delta}_i$ on z_i, to provide an estimate $\hat{\beta}$ of β and an estimate $\hat{\sigma}_\varepsilon$ of σ_ε.

iii) For each farm *i*=1,…n, the next four steps (a-d) are repeated B_1 times to yield a set of B_1 bootstrap estimates $\{\hat{\delta}_{i,b}* \;\; b = 1,...B_1\}$.

a) ε_i is drawn from the $N(0,\hat{\sigma}_\varepsilon{}^2)$ distribution with left-truncation at $(1-\hat{\beta}z_i)$.

b) $\delta_i* = \hat{\beta}z_i + \varepsilon_i$ is computed.

c) A pseudo data set (x_i*, y_i*) is constructed, where $x_i* = x_i$ and $y_i* = y_i\hat{\delta}_i / \delta_i*$.

d) A new DEA estimate $\hat{\delta}_i*$ is computed on the set of pseudo data (x_i*, y_i*), i.e. Y and X are respectively replaced by $Y* = \{y_i* \;\; i = 1,...n\}$ and $X* = \{x_i* \;\; i = 1,...n\}$ in program.

iv) For each farm i=1,…,n, the bias-corrected estimator $\hat{\hat{\delta}}_i$ is computed as follows:

$$\hat{\hat{\delta}}_i = \hat{\delta}_i - \widehat{\text{bias}}_i$$

where $\widehat{\text{bias}}_i$ is the bootstrap estimator of bias obtained as (Simar and Wilson, 1998):

$$\widehat{\text{bias}}_i = \left(\frac{1}{B_1}\sum_{b=1}^{B_1}\hat{\delta}_{i,b}*\right) - \hat{\delta}_i.$$

v) Maximum likelihood is used in the truncated regression of $\hat{\hat{\delta}}_i$ on z_i, to provide an estimate $\hat{\hat{\beta}}$ of β and an estimate $\hat{\hat{\sigma}}$ of σ_ε.

vi) The next three steps (a-c) are repeated B_2 times to yield a set of B_2 bootstrap estimates $\left\{(\hat{\hat{\beta}}_b*, \hat{\hat{\sigma}}_b*) \;\; b = 1,...,B_2\right\}$.

a) For each farm i=1,…n, ε_i is drawn from the $N(0,\hat{\hat{\sigma}})$ distribution with left truncation at $(1-\hat{\hat{\beta}}z_i)$.

b) For each farm i=1,…n, $\delta_i ** = \hat{\hat{\beta}}z_i + \varepsilon_i$ is computed.

c) Maximum likelihood is used in the truncated regression of $\delta_i **$ on z_i, to provide an estimate $\hat{\hat{\beta}}*$ of β and an estimate $\hat{\hat{\sigma}}*$ of σ_ε.

vii) Confidence intervals are constructed. The estimated $(1-\alpha)$ per cent confidence interval of the *j*-th element β_j of the vector β, is as follows:

$$\Pr\mathrm{ob}\left(Lower_{\alpha,j} \le \beta_j \le Upper_{\alpha,j}\right) = 1-\alpha$$

where $Lower_{\alpha,j}$ and $Upper_{\alpha,j}$ are calculated using the empirical intervals:

$$\mathrm{Prob}(-\hat{b}_\alpha \le \hat{\hat{\beta}}_j * - \hat{\hat{\beta}}_j \le -\hat{a}_\alpha) \approx 1-\alpha$$

where $Upper_{\alpha,j} = \hat{\hat{\beta}}_j + \hat{b}_\alpha$

$$Lower_{\alpha,j} = \hat{\hat{\beta}}_j + \hat{a}_\alpha .$$

The same method is applied to construct confidence intervals for the efficiency scores (Simar and Wilson, 2000).

CHAPTER II: THE DETERMINANTS OF THE TECHNICAL EFFICIENCY OF ACUTE INPATIENT CARE IN CANADA

Li Wang[1], Michel Grignon[2], Sheril Perry[3], Xi-Kuan Chen[3], Alison Ytsma[3], Sara Allin[3], and Katerina Gapanenko[3]

[1]Department of Health Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada
[2]Department of Economics, McMaster University, Hamilton, Ontario, Canada
[3]Canadian Institute for Health Information, Toronto, Ontario

Abstract

Objective: To evaluate the technical efficiency of acute inpatient care at the Pan-Canadian level and to explore the factors associated with inefficiency - why hospitals are not on their production frontier.

Data Sources/Study Setting: Canadian Management Information System (MIS) database (CMDB) and Discharge Abstract Database (DAD) for the fiscal year of 2012-2013.

Study Design: We use a non-parametric approach (Data Envelopment Analysis) applied to three peer groups (teaching, large and medium hospitals, focusing on their acute inpatient care only). The double bootstrap procedure (Simar and Wilson, 2007) is adopted in the regression.

Data Collection/Extraction Methods: Information on inpatient episodes of care (number and quality of outcomes) was extracted from the DAD. The cost of the inpatient care was extracted from the CMDB.

Principal Findings: On average, acute hospitals in Canada are operating at about 75% efficiency, and could thus potentially increase their level of outcomes (quantity and quality) by addressing inefficiencies. In some cases, such as for teaching hospitals, the factors significantly correlated with efficiency scores were not related to management but to the social composition of the caseload. In contrast, for large and medium non-teaching hospitals, efficiency related more to the ability to discharge patients to post-acute care facilities. The efficiency of medium hospitals is also positively related to treating more clinically non-complex patients.

Conclusions: The main drivers of efficiency of acute inpatient care vary by hospital peer groups. Thus the results provide different policy and managerial implications for teaching, large and medium hospitals to achieve efficiency gains.

Keyword: technical efficiency, acute inpatient care, peer group, Data Envelopment Analysis

Citation: Chapter II has been published in **Health Services Research Journal***:*

Wang L., Grignon M. Perry S., Chen X., Ytsma A., Allin S., and Gapanenko K. 2018, "The Determinants of the Technical Efficiency of Acute Inpatient Care in Canada", *Health Service Research,* DOI: 10.1111/1475-6773.12861

INTRODUCTION

Whether publicly or privately funded, hospitals in most health care systems are shielded from normal market competition for sound normative reasons. As a result, they may be technically inefficient. In a public single-payer system such as Canada (90% of hospital spending is publicly funded and most are not-for-profit), where, moreover, hospitals are funded mostly through global budgets rather than activity-based funding, the main concern is that resources allocated to hospitals are not put to best use, thus increasing wait times to access services or leading to overcrowding of hospitals (occupancy rates of Canadian hospitals are among the highest in the OECD). This is a crucial policy issue in Canada: hospitals use the issues of wait times and overcrowding to ask for substantial increases in funding (hospital spending fell from 45% of total health care spending in 1975 to 30% in the 2010s), while others use them to challenge the public monopoly on most procedures (several cases have been in discussion before provincial or federal Supreme courts). Provincial governments, who are responsible for funding and monitoring hospitals, want to keep rationing within a single-payer system and look for ways to decrease wait times, e.g., through shortening stays and reducing bed use for post-acute care (Baker et al. 2004, Sutherland et al. 2013), while keeping overcrowding and undesirable events such as nosocomial infections or mortality under control (Chan and Cochrane, 2016). In other words, they want to know more about how to increase outputs for a given level of inputs in the hospital sector, which amounts to measuring technical inefficiency and, even more importantly, understanding the policy-amenable levers that explains why some hospitals are not on the frontier.

This policy issue is the main motivation for this study: how can a rationed hospital sector find ways to admit more cases while achieving a good level of quality per case treated? Another motivation for this study is to contribute the broader discussion on the determinants of hospital efficiency: most of the international literature is concerned (and rightly so) with the role of ownership status (for profit versus not-for-profit or not-only-for-profit, e.g., Ferreira and Marques, 2015) and/or funding schemes (activity-based, global, or fee-for-service), but Canadian hospitals are homogenous along these two dimensions (provincial experiments with funding did not really start before 2012), and vary mostly in their environmental constraints such as size, population density or availability of post-acute home care or residential care, as described below. This study therefore explores two empirical questions: what is the overall level of technical efficiency in the production of acute inpatient care in Canada? And what factors drive inefficiencies within the hospital sector?

It is often misleading to monitor efficiency in one specific component of the health care system in isolation, precisely because health care is a system and all components are articulated together. As a result, one component might appear efficient simply due to an ability to shift costs to another component within the system. However, inpatient acute care provides highly specialized health care services, most of which cannot really be shifted to other providers. Also, Abelson et al., 2017 showed that access to timely care when needed is a key objective of the Canadian health care system and acute care hospitals play a crucial role in timely access and, therefore, in the overall efficiency of the system. Last, hospitals have more ability than primary care providers to devote resources to management and leadership.

Canadian studies published so far have either treated Canada as one data point in an international comparison or been restricted to one or two provinces. Varabyova and Schreyögg (2013) used OECD countries to identify a production frontier of inpatient care and concluded that Canada was more efficient, with fewer beds per capita and shorter average stays, than comparable countries. A study of Quebec hospitals found that up to 17% of costs could be saved through improved efficiency (Bilodeau et al., 2004). Asmild et al. (2013) assess the differences in hospital performance between Ontario and New Brunswick over the years 1998-2004 and suggest that efficiency could be improved by between 4% and 11% in Ontario compared with between 34% and 48% in New Brunswick.

No consensus as to what drives efficiency in the hospital sector has been reached, mostly due to methodological differences (Hollingsworth, 2008). The measurement of technical efficiency of hospitals is a cottage industry around the world and we rely mostly on systematic reviews conducted at regular intervals (Hollingsworth, 2003; Erlandsen, 2008; Hollingsworth & Peacock, 2008; Rosko & Mutter, 2008, O'Donnell and Nguyen, 2011, Giancotti & Mauro, 2015). A recent review of 200 studies conducted by O'Donnell and Nguyen (2011) identified a number of data problems that severely limit the validity of conclusions, specifically a lack of consistent definitions of costs and measures of quality. The same is true of Canadian studies of the drivers of efficiency. Appendix A1 describes the main findings of Canadian studies, which can be summarized as follows: hospital size, occupancy rate, case-mix index, urbanity and teaching status have significant influences on hospital efficiency (Gruca and Nath 2001; Bilodeau, Crémieux, and Ouellette 2000; Bilodeau et al., 2004; Bilodeau, Crémieux and Ouellette. 2009; Chowdhury, Wodchis and laport 2011; Asmid, Hollingsworth, and Birth 2013; Chowdhury and Zelenyuk 2016).

This study is the first to measure technical efficiency of Canadian hospitals from a Pan-Canadian perspective instead of within a single province. Even though hospital care is a provincial responsibility in Canada, it is organized along the same lines (public funding, global budget, not-for-profit ownership) in all provinces. We use a unique database of hospital characteristics at the national level (assembled by CIHI), and exploit the variability of contexts yet homogeneity in ownership status and funding models across Canada's provinces and regions to contribute to the discussion on the determinants of hospital efficiency.

DATA

Our unit of observation is the hospital and we use three data sources. Firstly, information on inpatient episodes were extracted from the Discharge Abstract Database (DAD, a standardized set of information on hospital stays provided by hospitals and harmonized and made accessible by CIHI) for the fiscal year 2012-2013. We aggregated the information on episodes by hospital, producing a number of cases measured by resource intensity weight (RIW). Secondly, we used the Canadian Management Information System (MIS) database (CMDB) for the same fiscal year to construct our input variable (inpatient cost). The CMDB is updated annually by provincial and territorial governments and then sent to CIHI for standardization and quality checks, and it contains financial information at the hospital level across Canada (CIHI, 2013). Lastly, we used publicly available hospital performance indicators (e.g., hospital standardized mortality ratio, readmission rates etc.).

Our main analysis is based on data for fiscal year 2012-13, but we also ran the analyses on data from fiscal year 2013-14 as a robustness check of our findings.

Selection and exclusion of hospitals

There are just over 700 hospitals in Canada, but we excluded all hospitals operating in the province of Quebec and Nunavut territory, because, due to differences in the way they collect their financial CMDB data, they do not have comparable input: the majority of hospital cost in these two jurisdictions doesn't cover the same perimeter as in the rest of the country. We therefore started with 576 hospitals.

We excluded 356 small hospitals because they do not really provide the same level of “care” as larger hospitals: most of them provide some surveillance and then transfer complex patients to larger hospitals (Asmild et al., 2013, Rechel et al 2016). The analysis of efficiency among small hospitals is the focus of a separate study. We also excluded another 41 hospitals that strongly concentrated on a specific type of care and do not offer the scope of treatments that most hospitals provide. We identified these hospitals using a Herfindal-Hirschmann index applied to the distribution of treatments provided by each hospital (OECD, 2011). Overall, we excluded 397 hospitals, leaving 179 hospitals in the analysis, of which 35 were teaching hospitals, 54 large hospitals and 90 medium hospitals. These 179 hospitals accounted for 84% of acute inpatient hospital costs and 94% of weighted cases discharged from hospitals in Canada outside of Quebec and Nunavut.

Peer groups

We ran all analyses by hospital peer groups (teaching, large, medium). Teaching hospitals are identified as those with confirmed teaching status from the provincial ministry or as teaching in the provincial ministry’s submission to the CMDB. Non-teaching hospitals are assigned to a large, medium or small hospital peer group based on their volumes and patient complexities, as described in Appendix A2 (CIHI, 2016d).

We conducted separate analyses by peer group to make sure that functionality and outputs are comparable across decision making units (Newhouse, 1994). Teaching hospitals are required to provide services that are not required of other hospitals. Large hospitals tend to provide a variety of highly specialized services that medium hospitals do not offer. Thus, a pooled frontier model estimated on all hospitals may be misleading (Gruca and Nath, 2001). Hospitals within each of the three peer groups have similar case mixes and service-distribution characteristics making it reasonable to assume they share a common production technology.

METHOD

Estimation of efficiency scores

Data Envelopment Analysis (DEA)

In this study, we use a non-parametric approach known as DEA to estimate a production frontier and the distance of each hospital to the frontier. DEA is one of the two main methods used to estimate hospital technical or cost efficiency (Chowdhury et al., 2016; Ferreira and Marques, 2016a; Varabyova Y. and Schreyögg J. 2013; Bilodeau et al., 2009). The main advantage of the DEA approach in the present case is that it allows for the estimation of frontiers for multi-outputs; in the case of inpatient acute care, we want to account for the quality as well as the quantity of stays a hospital produces (see below, sub-section "outputs"), because it reflects the dilemma governments face: how to reduce wait-times without jeopardizing quality?

We apply an output orientation version of the DEA approach which measures how more services could be produced with the same level of inputs. We assume variable returns to scale (increases in inputs do not result in proportional increases in outputs), which seems reasonable given diminishing returns in the hospital sector (Hollingsworth et al. 2003). A DEA on ratio data only would have to make the assumption of CRS, but, since we use here one scale-dependent ratio (number of stays) and one ratio (HSMR), VRS still applies (Hollingsworth and Smith, 2003).

There are two drawbacks of the DEA. First, DEA scores are serially correlated, i.e., the efficiency score for one hospital is not independent from the efficiency scores of other hospitals, which is problematic in the second stage regression analysis of the contributions of various factors to inefficiency (Bowlin,1998; Felder and Tauchmann, 2013). Second, it does not allow for a random component of efficiency. To address these drawbacks, we submit the initial DEA scores to a "smoothed bootstrap" method designed by Simar and Wilson (1998, 2007).

Outputs

Outputs are measured along two dimensions, quantity of stays and the quality of care. Quantity is measured as the number of discharges from acute care beds, weighed by the resource intensity weight (RIW) developed by CIHI, where each patient receives an RIW by the CMG+ methodology that groups patients by diagnoses, and relative resource use and adjusts for the presence of co-morbidities and interventions (CIHI, 2017). RIWs can be further adjusted for special cases (long stay outliers, transfers, death in the hospital or sign-out cases) called "atypical" weights (CIHI, 2013). However, in our analysis we use "typical" RIW which excludes the atypical cases because these adjustments provide resource for potentially inefficient care. As a robustness check, to ensure this choice did not drive our results, we run a sensitivity analysis using the normal RIWs which include the atypical weights.

We exclude hospitalizations and costs for mental health and rehabilitation, because they are not reported and weighted in a consistent way across provinces. Our measure of the quantity of outputs is based on inpatient discharges and, as a result, excludes day surgery and outpatient care. Outpatient care is considered separate from inpatient stays for the following reasons: first, most outpatient care is funded separately from hospitals' global budgets since specialists' payments flow directly from governments on a fee for service basis; second, outpatient care data is collected separately from inpatient care data and these data are only available for all hospitals in two provinces (Ontario and Alberta). Through this exclusion we are not able to assess the extent to which hospitals vary in their ability to substitute outpatient for inpatient care, and whether this affects the inpatient efficiency that we observe. However, we expect this exclusion to have a limited impact on the results given the limited ability for hospitals to substitute highly specialized acute inpatient care with outpatient care. More of a concern is the exclusion of day surgery as it can substitute for inpatient elective surgery care when a new, less invasive technology becomes available (e.g., percutaneous coronary interventions). If some hospitals substitute more than others (likely as a result of clinical decisions made differently by different physicians), they will be unfairly seen as less efficient on the quantity dimension, even though they produce the same number of treatments using fewer resources. Evaluation of the CIHI data shows almost 90% of interventions that should happen in day surgery are occurring in that setting - meaning most hospitals are consistently performing less invasive procedures in the day procedure setting). As a result, hospitals performing more day surgery will not be overly disadvantaged by the exclusion of day surgery from our measure of output.

Number of stays is not the only output of a hospital. Hospitals are not only supposed to treat as many patients as they can, given the resources made available to them, they are also expected to provide quality of care. Also, a trade-off between quantity and quality per stay is plausible, in the sense that resources allocated to quality improvement may prevent the hospital to admit more patients or that, in order to achieve a given level of quality, the hospital might keep their patients a bit longer but it doesn’t increase their RIWs (Valdmanis et al., 2008; Mckay and Deily, 2005). There is no positive measure of quality of inpatient care available in Canada (such as patient satisfaction) but negative events are recorded and can give an idea of the effort a hospital puts into treating its patients: a hospital will achieve better quality if it reduces its in-hospital mortality measured by Hospital Standardized Mortality Ratio (HSMR), its rate of nosocomial events (such as in-hospital sepsis), or the proportion of patients who have to be re-admitted after being discharged (CIHI, 2016c). Because DEA allows us to model several dimensions of output, we considered using several quality indicators as output, provided they pass three tests: (1) they were measured consistently across all hospitals in our sample, (2) they varied across hospitals (based on the coefficient of variation), and (3) the ratio of the variation in efficiency across hospitals due to the introduction of that particular dimension was not purely random (Badunenko et al. 2012). Mortality after major surgery failed the first test. Readmission rate failed the second test, as the coefficient of variation across hospitals was smaller than 10% (Nairy and Rao, 2003). In-hospital sepsis failed the third test, with a ratio of noise to total variance in efficiency of 92%. As a result, our only quality indicator for DEA available in the CIHI data source is the HSMR. Since mortality is an undesirable outcome, we transform HSMR by taking the inverse (1/HSMR*100) in the DEA so as to ensure that, all other things being equal, increased outputs should increase efficiency (Jacobs, Smith, and Street, 2006).

Inputs

There are three broad categories of inputs in the production of inpatient care (O'Donnell and Nguyen, 2011): capital investments (buildings, beds, etc.), labour (nurses, support staff, etc.) and other operating expenses (food service, etc.) Labour inputs have various skill levels and some hospitals rely more on labour and others on external providers, e.g. for food; to measure the various labour skill levels and types of resources we use dollar values of labour and other operating expenses aggregated into a broad category of operating costs. We exclude capital costs because they cannot be converted into yearly flows of expenditures. This latter exclusion is not too concerning because capital costs are small compared to operating expenses (CIHI, 2016a). In a robustness check analysis, we use total hospital expenditure (i.e., including capital investments) instead of inpatient only as our input variable to test that our methodological choice is not what drives the results.

Determination of factors explaining hospital efficiency

In the second stage of the analysis we treat the robust efficiency scores from bootstrap as dependent variables and used a bootstrapped truncated regression analysis to identify factors associated with these scores. These factors are characteristics of the hospital or of its environment, either actionable ones (the hospital and/or the regulator can do something about it) or descriptors of the unequal or diverse situations that hospitals face. Starting from a long list of potential factors based on the literature, we dropped those that were redundant with variables already considered in the first stage, inputs and outputs to the DEA (e.g. mortality after major surgery); those for which no measure was available in Canada (e.g. technology intensity); those that were not consistently measured (e.g. alternate level of care days). We were left with 15 potential determinants and their summary statistics and the underlying assumption in efficiency analysis are presented in A3 and A6 of the Appendix respectively.

Given the small number of observations (Draper and Smith, 1998), Principal Component Analysis (PCA) is applied to selected independent variables to summarize the dimensions of the hospital efficiency (Child, 1990). PCA is a data reduction technique which transforms the original set of variables into a smaller set of uncorrelated factors (built as linear combinations of the initial variables) that accounts for most of the variation in the original data set. At 0.7, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy is classed as great for PCA. Bartlett's test of sphericity $\chi^2=885.33$, $P < 0.001$, indicating the correlation between the explanatory variables was well defined for a PCA. A factor loading after oblimin rotation of .35 or above is used as an arbitrary value to designate a salient factor loading when interpreting the results of the exploratory analyses (Child, 1990). The PCA is performed on STATA 14.0.

Finally, we program the double bootstrap following Algorithm #2 in Simar and Wilson (2007) to estimate the marginal effect of the determinants of the hospital efficiency. For the steps 1-4 in Algorithm #2, we use the “dea”, “trunc.reg” and “rnorm.trunc” command in FEAR (Wilson, 2008) to calculate the bias-corrected efficiency score with the bootstrap results based on the truncated empirical normal distribution. Next, the steps 5 to 7 of Algorithm #2 are performed on STATA and consists in regressing the bias-corrected technical efficiency scores over a set of components derived by the PCA using a bootstrapped truncated regression in order to obtain unbiased coefficients and standard error.

RESULTS

Summary of data of hospitals, inputs and outputs

There were variations across peer groups in terms of total weighted RIW, HSMR and hospital expenditure on acute inpatient care (Table 1). There were approximately 32,000 weighted RIWs per hospital in teaching hospitals, 18,000 weighted RIWs in large and 5,000 in medium hospitals. The HSMR ranged from 45 to 150 across hospitals. The average expenditure on the acute inpatient care per hospital was $264.2 million in teaching hospitals, $100.5 million in large and $25.3 million in medium hospitals.

Table 1 insert here.

Robust efficiency scores

Table 2 summarizes the distribution of estimated robust efficiency scores. Average efficiency is 0.76 for all hospitals together with a standard deviation of 0.12 (range from 0.40 to 0.96): this means that hospitals could treat 24% more patients on average if all hospitals operated at maximum efficiency at their given input level, without increasing mortality per stay (hospitals could also decrease mortality per stay by 24% without treating more cases, but this would not address payers' concerns). Mean efficiency for teaching hospitals is 0.85, 0.78 for large hospitals and 0.72 in Medium hospitals.

Our results are robust to methodological choices or years of data:

When normal RIWs replace the typical RIWs, Kruskal-Wallis equality-of-populations rank test shows two sets of the technical efficiency scores of teaching hospitals are significantly different, but the difference for the large hospitals, $\chi2(1) = 1.139, p = 0.28$ or medium hospitals $\chi2(1) = 2.01, p = 0.16$. This indicates that the efficiency scores of teaching hospitals are sensitive to the atypical RIWs. It may be the fact that the teaching hospitals are more likely to treat more complex patients who may need very long stays. Using either total hospital cost or total inpatient cost instead of our measure of input generates qualitatively similar rankings of hospitals and the Kruskal-Wallis test also indicates that these excluded components do not have a significant effect on the efficiency and ranking. Last, performing the analysis on data from 2013 to 2014 generates change one can expect: hospitals do not stay at the exact same distance from the frontier but efficiency score rankings and distributions produce similar hospital rankings. Kruskal-Wallis equality-of-populations rank test shows there is no significant difference in the two sets of the technical efficiency, $\chi2(1) = 1.617, p = 0.20$.

Table 2 insert here.

The determinants of the technical efficiency

The PCA identifies four components representing 76% of total variance of the full hospital data. This is above the baseline parameter of 70% as recommended by Field (2009) and explains a significant amount of the original data. The first component is interpreted as a measure of advantaged hospital environments, because it is linked to more hospital beds, low proportion of rural patients and more percentage of individuals with post-secondary education in the region where the hospital is located. The second principal component strongly correlates with the proportion of long stay patients (negatively) and with maternity cases (positively) and it refers to a hospital with more non-complex patients. The third component is positively linked to the percentage transferred to long-term care and transferred to home with support services, and is interpreted as a measure of health system cooperation and availability of post- acute care. The forth component is linked to the percentage of elective surgery and having a major surgery (negatively), and overall readmission rate (positively), reflecting poor overall quality of treatment.

The double bootstrap truncated regressions on the teaching, large and medium hospitals show the following results: among teaching hospitals, highly efficient teaching hospitals are associated with advantaged environments. With respect to the large hospitals, increasing the proportion of non-complex patients is associated with improved hospital technical efficiency. Increasing cooperation within the health system and the availability of post- acute care beds is associated with improved efficiency for both large and medium hospitals. The medium hospitals, with advantaged environment, and higher proportion of non-complex patients are also more likely to achieve higher technical efficiency.

The regression models explain 52%, 65%, and 42% of variance in teaching, large and medium hospitals respectively, indicating substantial idiosyncratic variation in hospital management since our CIHI database allows us to capture most of the observable variation in hospital environment and management.

Table 3 insert here.

DISCUSSION

This study is the first to investigate the technical efficiency of acute inpatient care at the Pan-Canadian level. It contributes to the literature on assessing hospital efficiency by incorporating a quality of care metric into the specification of the output. Hospitals are not only supposed to process patients but also to restore health in the process. The number of discharges from acute care beds is the intermediate output of a hospital. The ultimate outcome of the hospital process is health improvement, which depends on the quality of the intermediate outputs. Some articles treat quality as a factor affecting efficiency rather than an output (Zuckerman et al., 1994; Jacobs et.al., 2012). Others estimate quality efficiency and quantity efficiency separately (Nayar and Ozcan, 2008). The former excludes quality of care from the production function and the latter ignores the trade-off between quantity and quality per stay (Valdmanis et al., 2008, Mckay and Deily, 2005). In our study, we included HSMR as an output in the production function. Other quality indicators, such as the overall readmission rate and in-hospital sepsis rate, enter the explanatory regression analysis on the hospital efficiency. This approach allows us to estimate potential gains in both quality and quantity of care when a hospital achieves its efficiency.

One strength is that we analyze efficiency by hospital peer groups. Each group has its production technology: teaching hospitals provide care for the most-complex patients and adopt new technologies and treatments sooner than non-teaching hospitals (Shahian et al.,2012) or, around 20% of medium hospitals in Canada do not perform major surgery, such as hip replacement. It is reasonable for a hospital to learn and improve from being benchmarked against the best performing hospitals from their own peers.

Another strength is that we considered a wide array of factors of efficiency, covering organization, management, demographics and patient characteristics, using PCA to summarize, instead of ad hoc selection or step-wise regressions (Grafen & Hails, 2002). Therefore, we can provide actionable recommendations to hospital and health system leaders to achieve greater efficiency. Specifically, among teaching hospitals, the social composition of the caseload (e.g., rural versus urban patients) is a key driver of efficiency, suggesting that provinces should re-distribute resources toward those teaching hospitals treating more rural patients. Simultaneously, teaching hospitals could receive help and resources to develop care plans across the continuum of heath care or provide remote care for patients from rural areas so as to be able to discharge them sooner. The main driver of efficiency among large and medium non-teaching hospitals is their ability to discharge patients to post-acute care facility; this again, suggests action from provinces, in increasing the number of post-acute care beds and personnel, but also on the ground to develop links between acute care hospitals and these facilities, as well as procedures to smooth the transition (e.g., discharges on weekends). Last, medium hospitals could improve their efficiency by transferring out more complex cases (to large and teaching hospitals) and concentrating on cases such as maternity.

Our conclusions are robust and indicative of the main determinants of efficiency among Canadian hospitals. We now briefly outline what remains to be done in this area of inquiry: First, we approach the output of hospitals by the number of discharges weighted by resources used. Ideally, however, the value of a hospital stay should be based on the value of the cure, rather than the resources put into it. The value of a cure is the difference between the health-related quality of life of the patient before and after the cure, and it does not always reflect the resources used in the treatment. One option to measure activity based on value for the patient is to use Quality Adjusted Life-Years (QALYs). However, this method raises several issues, mostly of practicality.

The main issue in using resources to measure value is that hospitals using more low-resource options are penalized. Statistics Canada weighted hospital activities based on their value to patients: they assigned the same RIW to both inpatient and outpatient activities within the same case mix group (Gu and Morin 2014). This is one option but it assumes implicitly that hospitals are always free to decide to go outpatient. If they are not, and can only outpatient milder cases, assigning the same RIW to outpatient and inpatient visits will penalize hospitals with more severe cases. The effect of outpatient care on the overall hospital efficiency is needed to be considered in the future.

Another limitation of our study is that HSMR is the sole indicator of quality of care. HSMR is an important component of the hospital output, but patient reported outcomes and patient reported experience would be interesting to add, once available.

Last, we excluded capital costs. Even though the sensitive analysis shows the capital cost doesn't influence the efficiency significantly, they might change efficiency at the margin. In addition, the exclusion of capital expenditures means that we do not distinguish between technologically innovative and more traditional hospitals and miss the effect of investment on efficiency (Valdmanis et al., 2008, Chowdhury et al., 2011).

CONCLUSION

The cross-sectional analysis of the hospital sector production efficiency in Canada using non-parametric DEA approach reveals that on average, in Canada, acute hospital care output could be increased by 24 percent with the same resources by eliminating inefficiency. Our analysis shows the following factors have positive influences on the technical efficiency of a hospital: more clinically non-complex patients, effective health system cooperation and availability of post-acute services, and advantaged hospital environments.

REFERENCES

Abelson, J., S. Allin, M. Grignon, D. Pasic, and M. Walli-Attaei. 2017. “Uncomfortable trade-offs: Canadian policy makers’ perspectives on setting objectives for their health systems.” *Health Policy* 121(1): 9-16.

Asmild, M., B. Hollingsworth, and S. Birch. 2013. “The scale efficiency of hospitals: One size does not fit all.” *Journal of Productivity Analysis* 40(2):197-206.

Badunenko, O., D.J. Henderson, and S.B. Kumbhakar. 2012. “When, where and how to perform efficiency estimation.” *Journal of the Royal Statistical Society*. 175(4): 863-892.

Baker, G.R., P.G. Norton, V. Flintoft, R. Blais, A. Brown, J. Cox, E. Etchells, W.A. Ghali, P. Hébert, S.R. Majumdar, and M. O'Beirne. 2004. “The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada.” *Canadian medical association journal* 170(11):1678-1686.

Bilodeau, D., P.Y. Cremiedx, P. Quellette, and T. Vovor. 2004. “Measuring hospital performance in the presence of quasi-fixed inputs: an analysis of Quebec hospitals.” *Journal of Productivity Analysis* 21(2):183-199.

Bilodeau, D., P.Y. Cremiedx, and P. Quellette. 2009. “Hospital performance in a non-competitive environment.” *Applied Economics* 41(4): 459-468.

Bilodeau, D., P.-Y. Crémieux, and P. Ouellette. 2000. “Hospital cost function in a non-market health care system.” *The Review of Economics and Statistics* 82(3): 489-498.

Bowlin, W.F.1998. “Measuring Performance: An Introduction to Data Envelopment Analysis (DEA).” *Journal of Cost Analysis* 15(2): 3-27.

Canadian Institute for Health Information (CIHI). 2013. *MIS Standards 2013*. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2016a. *National Health Expenditure Trends, 1975 to 2016*. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2016b. *Hospital Standardized Mortality Ratio: Technical Notes*. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2016c. *Health Indicators 2016*. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2016d. *Indicator Library: Peer Group Methodology*. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2017. *DAD Resource Intensity Weights and Expected length of Stay*. Ottawa, ON: CIHI.

Chan, B., and D. Cochrane. 2016. *Measuring Patient Harm in Canadian Hospitals. With What can be done to improve patient safety*? Canadian Institute for Health Information, Canadian Patient Safety Institute. Ottawa, ON: CIHI.

Child, D. 1990. *The essentials of factor analysis*. Cassell Educational.

Chowdhury, H., Wodchis, W., and Laporte, A., 2011. “Efficiency and technological change in health care services in Ontario: an application of Malmquist productivity index with bootstrapping.” *International Journal of Productivity and Performance Management*, 60(7): 721-745.

Chowdhury, H., and V. Zelenyuk. 2016. “Performance of hospital services in Ontario: DEA with truncated regression approach.” *Omega* 63: 111-122.

Draper, N.R., and H. Smith.1998. *Applied Regression Analysis,* pp.505-53. New York: Wiley – Interscience.

Erlandsen, E., 2008. “Improving the efficiency of health care spending.” *OECD Journal: Economic Studies* 1:1-33.

Felder, S., and H. Tauchmann. 2013. “Federal state differentials in the efficiency of health production in Germany: an artifact of spatial dependence?” *European Journal of Health Economics* 14(1): 21–39.

Ferreira, D.C. and R.C. Marques. 2015. “Did the corporatization of Portuguese hospitals significantly change their productivity?” *European Journal of Health Economics* 16(3): 289-303.

Ferreira, D.C. and R.C. Marques. 2016a. “Should inpatients be adjusted by their complexity and severity for efficiency assessment? Evidence from Portugal.” *Health Care Management Science* 16(1): 43-57

Field A. 2009. *Discovering Statistics using SPSS*, 3d Edition. London: Sage.

Giancotti, M., and M. Mauro. 2015. “Scale Efficiency of Public Hospital: A Content Analysis of 20 Years of Research.” *World Review of Business Research* 5(2): 24 – 38.

Ridic, G., Gleason, S. and Ridic, O., 2012. “Comparisons of health care systems in the United States, Germany and Canada.” *Materia socio-medica*, 24(2), p.112.

Grafen, A., and R. Hails. 2002. *Modern Statistics for the Life Sciences*. Oxford: Oxford University Press, Oxford.

Gruca, T. S., and D. Nath. 2001. “The technical efficiency of hospitals under a single payer system: the case of Ontario community hospitals.” *Health care management science* 4(2): 91-101.

Gu, W., and S. Morin. 2014. “Experimental measures of output and productivity in the Canadian hospital sector, 2002 to 2010.” *In Measuring Economic Sustainability and Progress*, pp. 575-594. University of Chicago Press.

Hollingsworth, B. 2003. “Non-Parametric and Parametric Applications Measuring Efficiency in Health Care.” *Health Care Management Science* 6(4): 203-18.

Hollingsworth, B., and S. J. Peacock. 2008. *Efficiency measurement in Health and Health care*. London, Routledge.

Hollingsworth, B & P. Smith. 2003. “The use of ratios in data envelopment analysis.” *Applied Economics Letters*, vol 10, no. 11, pp. 733-735.

Jacobs, R. 2001. Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis. *Health care management science* 4(2): 103-115.

Jacobs, R., P. C. Smith, and A. Street. 2006. *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*, Cambridge, UK: Cambridge University Press.

Jung, S. & Lee, S. 2011. “Exploratory factor analysis for small samples.” *Behavior Research methods* 43(3): 701-709.

McKay, N., and M. Deily. 2005. “Comparing High- and Low-Performing Hospitals Using Risk Adjusted Excess Mortality and Cost Inefficiency.” *Health Care Management Review* 30(4): 347–360.

Nairy, K.S., and K.N. Rao. 2003. “Tests of coefficients of variation of normal populations.” *Communications in Statistics Simulation and Computation* 32(3): 641-661.

Nayar, P., and Y.A. Ozcan. 2008. “Data Envelopment Analysis Comparison of Hospital Efficiency and Quality.” *Journal of Medical System* 32(3):193-199.

Newhouse, J.P. 1994. “Frontier Estimation: How Useful a Tool for Health Economics?” *Journal of Health Economics* 13(3): 317–322.

O’Donnell C and K. Nguyen. 2011. “Review of efficiency measurement methodologies to inform hospital resource allocation decisions in NSW: A Rapid Review.” Ultimo, NSW, Australia: Sax Institute.

OECD. 2011. “OECD Guidance, Competition Assessment Toolkit.” OECD Vol 2, Version 2.0.

Rechel, B., A. Džakula, G. Duran, N. Fattore, M. Edwards, M. Grignon, T. Haas, G. Habicht, P. Marchildon, A. Moreno, W. Ricciardi, L. Vaughan, and T. Anderson Smith. 2016. "Hospitals in rural or remote areas: a review of policies in 8 high-income countries." *Health Policy* 120(7): 758–769.

Rosko, M.D., and R.L. Mutter. 2008. "Stochastic frontier analysis of hospital inefficiency: a review of empirical issues and an assessment of robustness". *Medical Care Research and Review* 65(2):131-166.

Shahian, D.M, P. Nordberg, G.S. Meyer, B.B. Blanchfield, E.A. Mort, D.F. Torchiana, and S.L.T. Normand. 2012. "Contemporary performance of U.S. teaching and nonteaching hospitals." *Academic Medicine,* 87(6): 701–708.

Simar, L., and P.W. Wilson. 1998. "Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models." *Management Science* 44(1):49-61.

Simar, L., and P.W. Wilson. 2007. "Estimation and inference in two-stage, semi-parametric models of production processes." *Journal of Econometrics* 136(1): 31-64.

Sutherland, J. M., and R. T. Crump. 2013. "Alternative Level of Care: Canada's Hospital Beds, the Evidence and Options." *Healthcare Policy* 9(1), 26–34.

Valdmanis, V.G., M.D. Rosko, and R.L., Mutter. 2008. "Hospital quality, efficiency, and input slack differentials." *Health services research* 43(5p2):1830-1848.

Varabyova, Y., and J. Schreyögg. 2013. "International comparisons of the technical efficiency of the hospital sector: panel data analysis of OECD countries using parametric and non-parametric approaches." *Health policy* 112(1), 70-79.

Wilson, P. W. 2008. "FEAR 1.0: A Software Package for Frontier Efficiency Analysis with R." *Socio-Economic Planning Sciences* 42, 247--254.

Zuckerman, S., J. Hadley, and L. Iezzoni. 1994. "Measuring hospital efficiency with frontier cost functions." *Journal of Health Economics* 13(3): 255-280.

Table 1: Summary statistics of inputs and outputs by hospital peer group

	Teaching[I]		Large[II]		Medium[III]	
Input:	Mean	SD.	Mean	SD.	Mean	SD.
Inpatient Cost ex. mental and rehabilitation care (100,000')	2642	1502	1005	565	253	123
Output:						
Case mixed RIW (100')	384	197	162	105	35	22
Hospital Standardized Mortality Ratio	103	12	100	13	106	24

Notes:
SD is standard deviation
[I] Had confirmed Teaching status from the provincial ministry; or were identified as Teaching in the provincial ministry's submission to the Canadian MIS Database.
[II] Meet two of the following three criteria: ≥8,000 inpatient cases; ≥10,000 weighted cases; ≥50,000 inpatient days.
[III] Medium: Do not meet large community hospital criteria and ≥2,000 weighted cases.

Table 2: Summary statistics of bias-corrected efficiency scores[IV] by hospital peer group

Hospital Peer Group	N	Mean	SD	Range	
				Min.	Max.
Teaching[I]	35	0.846	0.085	0.652	0.962
Large[II]	54	0.785	0.106	0.57	0.961
Medium[III]	90	0.717	0.127	0.405	0.938
All hospital	179	0.763	0.123	0.405	0.962

Notes:
SD is standard deviation
[I] Had confirmed Teaching status from the provincial ministry; or were identified as Teaching in the provincial ministry's submission to the Canadian MIS Database.
[II] Meet two of the following three criteria: ≥8,000 inpatient cases; ≥10,000 weighted cases; ≥50,000 inpatient days.
[III] Medium: Do not meet large community hospital criteria and ≥2,000 weighted cases.
[IV] The estimation of the DEA efficiency scores using smoothed bootstrap" method by Simar and Wilson (1998, 2007).

Table 3: Determinants of technical efficiency (double bootstrap regression) using the components from PCA

Component	Description of the components	Coefficient	Bootstrap Std. Error
Teaching Hospital[I]			
Component 1	Advantaged environments	0.047***	0.011
Component 2	Higher proportion of non-complex patients	0.013	0.01
Component 3	Health system cooperation and availability of post- acute care	0.008	0.017
Component 4	Poor overall quality of care	0.014	0.012
R^2		52%	
Large Hospital[II]			
Component 1	Advantaged environments	0.027	0.017
Component 2	High proportion of non-complex patients	0.036**	0.009
Component 3	Health system cooperation and availability of post- acute care	0.03***	0.01
Component 4	Poor overall quality of care	0.003	0.009
R^2		65%	
Medium Hospital[III]			
Component 1	Advantaged environments	0.03***	0.01
Component 2	Higher proportion of non-complex patients	0.022***	0.007
Component 3	Health system cooperation and availability of post- acute care	0.016**	0.007
Component 4	Poor overall quality of care	0.03	0.031
R^2		42%	

Note: *** p<0.01, ** p<0.05, * p<0.1

[I] Had confirmed Teaching status from the provincial ministry; or were identified as Teaching in the provincial ministry's submission to the Canadian MIS Database.

[II] Meet two of the following three criteria: ≥8,000 inpatient cases; ≥10,000 weighted cases; ≥50,000 inpatient days.

[III] Medium: Do not meet large community hospital criteria and ≥2,000 weighted cases.

Appendix:

A1: Selected Canadian studies in the hospital efficiency

Authors	**Region**	**Covered year**	**# of hospitals**	**Method**	**Results**
Bilodeau et al.,(2004)	QC	1981-82 through 1992-1993	121 short-term hospitals	DEA/RTS	QC financial DRG; Inpatient diversity index; inpatient complexity index; outpatient diversity index; outpatient complexity index; % of patients over 65; increases in patients over 65 since 1981
Bilodeau et al., (2009)	QC	1981-82 through 1992-1994	121 short-term hospitals	SFA-translog (cost function)	the size of the hospital, the average age of patients, the teaching status, the socio-demographic characteristics of the population, the variety of services , the level of technological changes
Bilodeau et al., (2000a)	QC	1981-1982 through 1992-1993	121 short-term hospitals	translog cost function	The properties of the non-market hospitals' cost functions are compatible with short-term, but not long-term, cost-minimizing behavior.
Gruca et al.,(2001)	ON	1986	168 community general hospitals	DEA(pooling vs. nesting with and without LTC beds);	There is no significant differences in efficiency across ownership (government, religious or secular non-profit); Nor do we find significant differences in efficiency by size or location.
Chowdhury et al.,(2011)	ON	2002/2003 to 2005/2006	113 acute hospitals	DEA, bootstrapping Malmquist Productivity Indices	A large number of hospitals did not achieve significant progress in terms of productivity. By taking geometric means of estimates for all years it was observed that while overall productivity and efficiency of hospital in Ontario declined during the study period, technological progress increased at a rate of 5.95 percent on average.

Asmild et al., (2013)	ON, NS	1998-2004	117 (ON), 25(NB) public hospitals	DEA, Malmquist indices	Efficient hospitals in the larger more urban province are larger than efficient hospitals in the smaller more rural province.
Chowdhury et al.,(2016)	ON	2003-2006	113 acute hospitals	DEA, bootstrapping	The occupancy rate, rate of unit-producing personnel, outpatient–inpatient ratio, case-mix index, geographic locations, size and teaching status are significant determinants of efficiency.

A2: Hospital peer groups

Teaching
• Had confirmed Teaching status from the provincial ministry; or
• Were identified as Teaching in the provincial ministry's submission to the Canadian MIS Database.
Community — Large
2 of the following 3 criteria:
• ≥8,000 inpatient cases
• ≥10,000 weighted cases
• ≥50,000 inpatient days
Community — Medium
Do not meet large community hospital criteria and ≥2,000 weighted cases
Community — Small
Do not meet large community hospital criteria and <2,000 weighted cases

Source: Canadian Institute for Health Information. Indicator Library: Peer Group Methodology. Ottawa, ON: CIHI; 2016.

A3: **Glossary of Terms**

HSMR	The ratio of the actual number of in-hospital deaths in a region or hospital to the number that would have been expected, based on the types of patients a region or hospital treats. An HSMR equal to 100 suggests that there is no difference between the hospital's mortality rate and the overall average rate; greater than 100 suggests that the local mortality rate is higher than the overall average; and less than 100 suggests that the local mortality rate is lower than the overall average.
% Long stay patients	The proportion of hospitalizations with total length of stay greater than trim point defined by case mix group.
% Acute patients transfers (in or out)	The percentage of patients who had been transferred from or to another acute hospital.
Beds (log form)	The number of acute beds (log form)

Nursing hours per weighted case	The number of worked hours from all personnel (excluding medical personnel) in hospital nursing units to produce a weighted case. Numerator: The numerator includes all worked and purchased hours in nursing inpatient functional centres. The numerator is adjusted for the proportion of inpatient activity determined by workload/activity statistics. Denominator: The denominator includes total acute inpatient weighted cases (obtained from the DAD), excluding day procedures.
% Rural patients	Proportion of patients who resided in the rural area
% Neighbourhood post-secondary education	The percentage of individuals with post-secondary education in the region where the hospital is located
% Transferred to long term care for first time	The percentage of patients who transferred to long term care at the first time after acute inpatient care.
% Overall readmission rate	the risk-adjusted rate of urgent readmissions within 30 days of inpatient discharge for the following patient groups: obstetric, patients age 19 and younger, surgical and medical.
% Discharged home with supportive care	The percentage of patients discharged home with supportive care after acute inpatient care.
% Expenditures on administration	The percentage of the legal entity's expenses that were spent in administrative departments, such as finance and human resources, to total expense. Numerator: The numerator includes all expenses associated with the administrative, finance, human resources and communication functional centres. Denominator: The denominator includes all expenses net of recoveries.
% Elective surgery	Surgery that is scheduled in advance because it does not involve a medical emergency
% In-hospital sepsis	The risk-adjusted rate of sepsis that is identified after admission
% Newborns	The proportion of the new born cases

A4. Summary Statistics of the potential determinants of the hospital efficiency by hospital peer group

Variables	**Teaching**[I]		**Large**[II]		**Medium**[III]		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
% Long stay patients	4.3	1.2	4.5	2.6	5.8	2.7	5.1	2.5
% Acute patients transfers (in or out)	9.1	3.9	5.9	2.4	11.1	4.3	9.1	4.3
Beds (log form)	6.3	0.5	5.8	0.4	4.5	0.5	5.2	0.9
Nursing hours per weighted case	43.8	6.1	42.7	10.9	48.1	13.6	45.6	11.9
% Rural patients	15.1	10.9	15.5	15.1	37.1	22.9	26.2	21.7
% Neighbourhood post-secondary	78.8	2.8	77.6	4.3	75.3	5.2	76.7	4.8
% Transferred to long term care for first time	1.4	0.9	2.1	1.1	2.2	1.4	2	1.3

% Overall readmission rate	9.1	0.8	8.6	0.8	8.8	1	8.8	0.9
% Discharged home with supportive	10.6	6.4	9.8	5.4	11.2	8.1	10.7	7.1
% Expenditures on administration	4.8	2.2	5	1.6	5.6	1.9	5.3	1.9
% Elective surgery	28.6	6.6	22.7	5.4	15.2	7.7	20.1	8.7
% In-hospital sepsis	4.6	2	4	1.9	3	1.8	3.6	2
% Newborns	12	7.4	12.7	5.4	10.3	5.6	11.3	6
Hospitals performs major surgery	1	0	1	0	0.8	0.4	0.9	0.3
Hospital has a maternity department	0.9	0.4	0.9	0.2	0.9	0.3	0.9	0.3

[I] Had confirmed Teaching status from the provincial ministry; or were identified as Teaching in the provincial ministry's submission to the Canadian MIS Database.

[II] Meet two of the following three criteria: ≥8,000 inpatient cases; ≥10,000 weighted cases; ≥50,000 inpatient days.

[III] Medium: Do not meet large community hospital criteria and ≥2,000 weighted cases

A5. Rotate factor loading from the PCA by hospital peer group

	Comp1	Comp2	Comp3	Comp4
% Long stay patients		-0.462		
% Acute patients transfers (in or out)		-0.3536		
Nursing hours per weighted case			-0.3941	
number of beds (log form)	0.4917			
% Rural patients	-0.4454			
% Neighbourhood post-secondary education	0.354			
% Transferred to long term care for first time			0.5572	
% Discharged home with supportive care			0.5295	
% Expenditures on administration		0.3937		
% Elective surgery				-0.3916
% Overall readmission rate				0.7969
% In-hospital sepsis				
Hospitals performs major surgery (yes/no)				-0.4565
% Newborn		0.5064		
Having a maternity department (yes/no)		0.4833		

Note: Oblimin rotation used and the factor loadings more than 0.35 are retained.

A6: **The underlying assumption of each potential determinant in efficiency analysis**

- % Long stay patients: hospitals that have more stays identified as “long stays” can be less efficient in their ability to discharge patients.
- % Patient transferred (in or out): transfers may help patients to obtain more effective treatment and make for better quality patient care.
- % Patients discharged to home with care: hospitals that can discharge patients with homecare are more efficient as they can discharge patients earlier than other hospitals.
- Transferred to long term care for first time: hospitals that can discharge patients to long-term care institutions are expected to be more efficient as they can discharge patients earlier than other hospitals.
- Nursing hours per weighted case: nursing costs are the largest portion of hospital budgets (excluding physician remuneration). Hospitals that manage nursing labor better are expected to be more efficient.
- % Expenditures on Administration: hospitals that invest more in back office health services management could be more efficient in the way they use frontline resources.
- In-hospital sepsis: hospitals with higher in-hospital sepsis tend to be less efficient.
- 30-day overall readmission: hospitals with higher overall readmission tend to be less efficient.
- % patients living in rural areas: Hospitals tend to keep those patients longer because lack of services in the community.
- Level of education in the region where the hospital is located: It is expected that hospitals with a higher portion of educated patients will be healthier, allowing for the hospital to be more efficient.
- Number of Beds: hospitals with more beds could be more efficiency due to the economic scale.
- % Elective surgery: Elective Surgery is planned, and outcomes are usually good and thus might be less expensive than other types of admission
- % new born: normal Newborns/Delivery's are non-complex and not costly cases, hence hospitals with more maternity cases might be more efficient.

CHAPTER III: SPATIAL INTERDEPENDENCE OF EFFICIENCY OF ACUTE INPATIENT CARE IN CANADA

ABSTRACT

Objective. To evaluate the interdependence of productive efficiency across hospitals in Canada: is efficiency in producing acute inpatient care in one hospital affected by the efficiency of neighbouring hospitals? There can be negative influence, if hospitals are in competition (e.g., shift higher risks to other hospitals) or positive ones if hospitals collaborate and transfer patients according to need.

Data Sources/Study Setting. We use administrative data collected by the Canadian Institute for Health Information (CIHI): The Canadian Management Information System (MIS) database (CMDB), the Discharge Abstract Database (DAD) for fiscal year 2012– 2013, and addresses of hospitals from Hospital Service Providers (HSP).

Data Collection/Extraction Methods. Information on acute inpatient episodes of care (number weighted by intensity of care and quality of outcomes) was extracted from the DAD. The cost of the inpatient care was extracted from the CMDB. Spatial contiguity (distance) and binary (province) weighting matrix are derived using GIS. Determinants of efficiency scores (such as average length of stay or level of education in the population living in the catchment area of the hospital) were extracted from Indicators of Health System Performance (a database on hospital maintained by CIHI).

Study Design. The first stage is to generate standard bias-corrected efficiency scores through data envelopment analysis. The second stage is to detect spatial interdependence on these efficiency scores using the spatial regression model (SRM), varying the distance threshold used to define maximum distance of influence. We then estimate the influence of determinants of efficiency controlling for the effect of spatial interdependence among hospitals.

Principal Findings. The analysis shows substantial and significantly positive spatial interdependence effects on the efficiency of acute inpatient care in Canada: with maximum distance defined as120km, productive efficiency in acute inpatient care in one hospital is increased by 0.3% for each percentage increase in efficiency of other hospitals in the area. These are robust to changes in the definition of catchment areas but there is no spatial interdependence effect if we use provincial boundaries to the definition of maximum distance for influence. However, adding spatial interdependence to the list of determinants of productive efficiency does not affect their estimated influences and we find that hospitals with fewer long term patients, more acute transfers and more beds are more efficient.

Conclusion. Hospitals in Canada chose cooperative strategies rather than competition to improve their technical efficiency, through transfers of patients on the basis of need. This may be linked to the main mode of payment (global budget), even though many provinces are now moving toward activity-based funding.

Key Words. Interdependence, Technical efficiency, Data Envelopment Analysis, Spatial regression model, Spatial weighting matrix

Highlights:

- It is the first study to examine the spatial interdependence of the efficiency of acute inpatient care in Canada, and one of the few studies to do so internationally.
- There is substantial and significantly positive spatial spillover on the technical efficiency of the acute inpatient care in Canada.
- By taking interdependence into account, the main determinants of the productive efficiency of acute inpatient care are the number of long term patient (negative effect), acute transfers (positive effect) and acute beds (positive effect).

1. Introduction

Hospitals pose an interesting challenge to analysts of productive (technical and cost) efficiency. The standard model of efficiency analysis posits that decision making units (DMUs) work in isolation to produce the same outputs, following the same technology, and in a somewhat competitive environment: the resources they use are made available to them on a competitive market that sets their unit costs and they never benefit from the outputs produced by other DMUs. Hospitals, however, are different: they certainly compete with each other to some extent, trying to attract patients, private donations and clinicians, but they also sometimes complement each other and work in networks of care: if a hospital cannot treat a given patient, it must transfer them to another hospital with whom they have a stable (reciprocal or hierarchical) relationship. Moreover, not-only-for-profit hospitals, who provide the vast majority of treatments in all health care systems (even in the US) do not (and cannot) have their bottom line as their main objective, but rather timely access to quality care for the community they serve; to reach this objective, they work in collaboration with nearby hospitals.

One specific instance of collaboration is hierarchical, i.e., collaboration between hospitals at various levels of specialization. When such collaboration takes place, more than one hospital to treat a patient and contribute to one episode of care, and the resources and technologies used by each hospital are often vastly different. This type of collaboration is at play in most transfers to teaching and large hospitals that have specialized care (mental health, pediatric care, women's health, heart etc.). In some cases, one large hospital in a region is responsible to perform almost all procedures in one specialty, such as angiography or percutaneous cardiac intervention (PCI). Large hospitals can sit at both ends of large volumes of transfers receiving patients in for the type of treatments they are specialized in, and sending patients out for treatments for which another hospital is better positioned or to local hospitals for rehabilitation or surveillance once specialty care is no longer needed.

It is well established that hospitals collaborate, and it is theoretically plausible that collaboration might lead to better performance. It is also supported empirically: previous analysis of efficiency of hospitals in Canada (conducted by a team including these authors) found the volume of transfers (in and out) to be an important determinant of efficiency, in the sense that a hospital with more transfers is more efficient (more treatments and lower risk per treatment, Wang et al., 2018). Analyses of efficiency of hospitals should always try to control for the levels of efficiency of hospitals susceptible of interacting with each other as a factor explaining performance.

Beside collaboration (transfers), there are other reasons why hospitals should be treated as inter-dependent DMUs rather than isolated DMUs with no influence on each other.

There are reasons to believe that hospitals within the same catchment area will share crucial environmental characteristics that will in turn lead to a common component of their level of efficiency (positive interdependence, see for instance, Mobley 2003; Mobley et al., 2008): firstly, they share the same patient pool and, if these pools differ across catchment areas along dimensions that are not captured by case-mix adjustments, hospitals in the same area will share some unobservable heterogeneity. Secondly, physicians practicing in the same area may share practice styles (e.g., because they are more likely to have received their training at the same school or because they communicate with others), or the same physician may practice in several hospitals in the same area, and this common style will create for all hospitals in that area a common unobserved (and likely non observable) factor of efficiency[2] (Herwartz & Strumann 2010). Thirdly hospitals operate in a regulated environment where prices are set by a central payer (rather than determined by the market as the result of price-competition among providers), and, as a result, hospitals can only compete by comparison with their peers, again reinforcing positive interdependence of levels of performance within a given geographic area (being of better quality than their comparators, Gravelle et al., 2014; Longo et al., 2017). The results of the UK study (Gravelle et al., 2014) show that an increase in rivals' quality by 10% increases a hospital's quality by 1.7%–2.9% with hospital markets defined by catchment areas of 30-minute driving time. Lisi et al., (2017) also find that hospital dimensions of quality are strategic complements either in the local hospital market or outside their local market in Italy. Gaynor et al., (2016), using a difference-in difference approach, find that higher density of hospitals leads to better health outcomes, a finding they interpret as a positive effect of competition.

On the other hand, the performance of two hospitals in the same catchment could be negatively correlated. Under global budget or activity-based funding, hospitals have an incentive to cherry-pick patients and shift the cost of more complex, vulnerable and costly patients (either relative to the community or within a given diagnosis-related group) to other hospitals (Norton et al., 2002). Whether positive or negative, interdependence means that the performance of one hospital affects the performance of its neighbors. Ignoring the dependence among hospitals would lead to bias in the estimation of efficiency levels (and ranking) as well as possibly erroneous conclusions on the significance of the determinants of such efficiency levels. It has policy implications, too:

---

[2] In Canada, physicians are in private practice and bill for their services independently of the hospital. Their costs are not part of the hospital's budget, but their activity contributes to and shapes to a large extent the quantity and quality of what the hospital produces. Hospitals do not have much control on physicians practice, but their performance depend on it (Currie et al. 2016; Pai et. al 2000).

for instance, introducing competition among hospitals might decrease the incentive to collaborate and lead, paradoxically, to lower quality levels on average and a waste of social resources. But it may also increase quality if hospitals compare to each other and know that patients use quality metrics to choose among hospitals. In both instances, interdependence among rivals or partner's matters (Gaynor et al., 2016; Propper 2012; Cooper et al., 2011).

Economic geography (or quantitative spatial economics, Redding and Rossi-Hansberg, 2017) has developed tools to study and measure interactions among agents distributed in space or positions in a network and the general idea of assessing the role of interdependence in efficiency is to merge the literature on efficiency (mostly Data Envelopment Analysis in the case of hospitals, see Hollingsworth, 2008) and the econometric spatial regression model. It solves the problem of interdependence by taking into account observed and unobserved linkages of hospitals in the same catchment area (Getis et al. 2004; Srinivasan 2008). It generates a parameter called spatial lag that characterizes the degree of strategic complementarity or substitutability among peer hospitals. A key assumption in the development of such regressions is the construction of hospital markets and their representation by a spatial weight matrix (Gravelle et al., 2014; Mobley et al., 2008). Dormann et al., (2007) proposes a method, the Conditional Auto Regressive (CAR) model, that allows the analyst to extract the spatial weights matrix from the data, without having to make a priori assumptions.

The present analysis of acute inpatient care in general hospitals in Canada shows significant positive interdependence (spillover) across hospitals and their levels of efficiency: an increase in its "neighbors" efficiency by 10% increases the efficiency of a given hospital by approximately 2-3%, depending on the type of spatial weight matrix used. Our study contributes to an emerging discussion on the role of interdependence, peer influence, and shared environments in the performance of hospitals (and DMUs more broadly). To the best of our knowledge, there are only three other such analyses in the hospital sector. One, based on German data for 2002-2006, finds both positive correlation (environmental effect) and negative spatial spillovers among hospitals, the introduction of activity-based funding reinforcing the latter to the detriment of the former, something they describe as cost-shifting motivated by hidden costs to some hospitals (Herwartz & Strumann 2010). Second, based on the American Hospital Association (AHA)'s Annual Survey of Hospital for FY 1997, Ferrier and Valdmanis (2006) found that a hospital is more efficient the greater the average efficiency of its peers. A more recent one, based on British data from 2010 to 2013 cannot identify any significant interaction effect on efficiency, but it uses a very specific set of efficiency measures (occupancy rate and cancelled operations, Longo et al., 2017). We discuss below the reasons why Canadian hospitals might show positive spillovers in efficiency when German ones show negative, and British ones no spillover at all.

2. Method

The modeling framework used for this study is shown in Figure 1. The first stage is to generate standard bias-corrected efficiency scores through DEA. The second stage is to detect spatial interdependence in these efficiency scores, using GIS coding of addresses of hospital to calculate hospital to hospital distances and create catchment areas around each hospital; the parameter of spatial interdependence is then entered in the regression explaining efficiency scores by standard determinants of efficiency such as average length of stay, number of beds or the education level of the population living in the hospital's catchment area, categorized as environmental, organizational and managerial factors.

2.1 Data

The data to estimate efficiency scores are drawn from two distinct sources. Information on inpatient episodes of care are extracted from the Acute Care Discharge Abstract Database (DAD), a standardized set of information provided by provinces and harmonized and made accessible by CIHI for the fiscal year 2012-2013 (April 2012 to March 2013). The other data source is the financial data from the Canadian Management Information System (MIS) database (CMDB) (CIHI, 2013) for the fiscal year 2012-2013. CMDB consists of financial and statistical information on Canadian hospitals, updated annually from data submitted by provincial and territorial ministries.

The population of hospitals included in this study is comprised of large, teaching and medium hospitals in all provinces except Quebec, because the data in that province are not comparable to those from the rest of Canada. Small or rural hospitals which are defined as less than 2,000 weighted inpatient cases per year on the basis of the intensity of resources (CIHI, 2016a; Asmild et al. 2013) are also excluded because they seem to serve a different purpose in the provincial health care system than other hospitals. Small hospitals often act essentially as community health centers that support the triage of acute patients as well as the post-acute recovery of patient discharged from larger hospitals. Small hospitals are sometimes also kept open in remote areas for political reasons (they are the main employer in the area) and, as a result, their production function is not comparable to that of other hospitals and it must be treated separately. Our selection criteria result in a total of 179 hospitals, of which 35 are teaching, 54 large and 90 medium hospitals.

2.2 Estimation of the efficiency score of acute inpatient care

As is standard in that literature, we use Data Envelopment Analysis (DEA) to estimate the scores. It is the preferred method because it is non parametric and there is no obvious production function of acute inpatient care that can be derived theoretically (hence the specification of the assumptions of a stochastic frontier analysis would be hard to justify). We estimate output-oriented efficiency scores because hospitals are under global budget with an objective to maximize health or quality-weighted treatments given their resources. We assume variable returns to scale because of evident diminishing marginal returns (adding more resources increases output by less) when one examines the relationship between spending on aggregate and health outcomes (OECD, 2011).

Another advantage of the DEA approach is that it allows the analyst to consider multi-output production functions: a hospital does not only treat as many patients as possible, it is also supposed to provide quality of treatment (often measured as the absence of adverse events such as mortality, re-admission, or sepsis): based on data availability, we characterize the output of acute inpatient care units in hospitals as comprised of the number of treatments weighted by resource intensity (a relative scale developed by CIHI (2017) to characterize the amount of resources used in a given treatment) and in-hospital mortality (as our measure of quality of treatment) (CIHI, 2016b). The expenses on the acute inpatient care is used as the single input (Wang et al. 2018). Following Simar and Wilson (1998, 2007), we apply a smoothed bootstrap DEA method to introduce some random error into the DEA estimate and to account for serial correlations of efficiency scores: this approach reduces the over-optimistic bias inherent in the DEA method and, most importantly, purges efficiency scores of their serial correlation, a requirement to use these scores in a second stage regression explaining them by characteristics of hospitals (see Wang et al. 2018 for more detail).

The Data Envelopment Analysis (DEA) is applied with FEAR package in Rstudio software (Wilson, 2008).

2.3 Spatial Interdependence

Interdependence can be defined as the relationship between proximity and the variable of interest (here, efficiency): do hospitals that are closer (according to a definition of proximity) influence each other's efficiency? The first methodological decision in defining interdependence is therefore to choose a measure of proximity. In this study, we define proximity as geographic proximity (distance between two hospitals, see Appendix A1), following Tobler's first law of geography (Tobler 1970) that everything is related to everything else, but near things are more related than distant things. The underlying idea is that hospitals collaborate and compete through patient flows (either transfers, capture or shifting) and these flows are determined in large measure by geographic distance (it is easy to transfer when distance is not too large or travel time is not too long). We acknowledge that other definitions of proximity could be used: for instance, two hospitals can be geographically close but politically distance if they sit across a jurisdictional border (different provinces in Canada) and a hospital might influence a more distant hospital located in the same province than a closer hospital across the border. It could also be the case that distance is not the same as travel time, due to geographic obstacles (mountain range or water), or that hospitals develop networks with distant hospitals for idiosyncratic reasons (e.g., reputation or personal links between physicians). We make the assumption that distance is the main determinant of proximity and intensity of relationship between hospitals, run a sensitivity analysis using political distance as a definition of proximity, and consider that travel time or networks are second order determinants only.

2.3.1 Spatial Weighting Matrix

Once we have a definition of proximity, we calculate a matrix of proximity weights, called the Spatial Weighting Matrix. These weights are used in the calculation of the statistic for the spatial autocorrelation test and in the estimation of the SRM proper.

In the geographic distance definition of proximity (our preferred definition), spatial weights reflect the fact that two distant hospitals are not correlated (small weight in the cell corresponding to these two hospitals) and two close hospitals are highly correlated (large weight). Weights are therefore a negative power of distance. It is also standard in this area to set a threshold beyond which correlation is set at zero whatever the actual weight (if two hospitals are too distant, their correlation is assumed to be negligible). Last, weights are standardized so that they are always unit-less and relative. This supposes three methodological choices: a value for the power parameter, a value for the threshold, and whether to standardize across rows or columns. We choose a row-standardization (it allows us to treat any product of the matrix by a vector of characteristic of hospitals to yield the vector of distance weighted average values of the characteristic for all hospitals, the distances in row *i* being calculated relative to hospital *j*), with a threshold of 60, 120, 180 or 240 kilometers (four variants) and a power of 1.31. The value for the power comes from a study by Shahid et al. (2009) who found that the best approximation of proximities based on road distance or travel time is provided by proximity measured as geographic distance with that parameter. The matrix terms are therefore as follows:

$$W_{ij} = \begin{cases} \dfrac{d_{ij}^{-\delta}}{\sum_{j=1}^{n} d_{ij}^{-\delta}}, if\, d_{ij} < d, i \neq j, \delta > 0 \\ 0, otherwise; \end{cases}$$

with d_{ij} is the distance between hospital *i* and *j*, δ=1.3 the distance decay parameter, and *d* is the threshold distance (spillovers are assumed to fall to 0 beyond the threshold).

In the political definition of proximity (variant), the matrix is binary in the sense that each cell takes a value of 1 if two hospitals are in the same province and 0 otherwise.

2.3.2 Autocorrelation Test

Based on the weights, Moran's I statistic (Moran, 1948, 1950a, 1950b) can be calculated as:

$$I = \frac{N}{W} * \frac{\sum_i \sum_j w_{ij}(v_i - \bar{v})(v_j - \bar{v})}{\sum_i (v_i - \bar{v})^2}$$

where *N* is the number of spatial units indexed by *i* and j, w_{ij} is the matrix of spatial weights, v_i is the vector of efficiency scores, *W* is the sum of w_{ij} and $\bar{v}$ is the mean of the efficiency score. Moran's *I* statistics takes values in the (-1, 1) range, with 1 indicating perfect spatial similarity (efficiency scores correlate positively among hospitals that are geographically close to each other), 0 indicating no autocorrelation, and -1 indicating perfect dispersion (negative correlation between efficiency and spatial proximity).

2.4 Spatial Regression Model

If spatial autocorrelation is tested to be significant, based on Moran's *I* statistic, the next and final step is to use the Spatial Regression Modelling (SRM) (Cliff and Ord, 1981; Anselin 1988b) to correct for the spatial autocorrelation error in a regression model of efficiency scores on determinants of efficiency. The OLS estimates of the coefficient are potentially biased when proximity (spatial autocorrelation) is not controlled for. When the SRM controls for spatial autocorrelation, two variants of the SRM exist: the Spatial Lag Model (SLM) simply adds the distance-weighted average efficiency of all neighbouring hospitals to the OLS regression of hospital efficiency scores:

$$\theta = \rho W\theta + \beta X + \epsilon, \, \varepsilon N(0, \sigma^2 I)$$

Where θ is the $N \times 1$ vector of observations (bootstrapped efficiency scores estimated in the first stage), X is the $N \times K$ matrix of observations of K explanatory variables (determinants of efficiency, such as: size, type of hospital, case-mix etc., determined in a previous paper, Wang et al. 2018), W is the matrix of spatial weights (see 2.3.1). The spatial lag parameter ρ (which is the main parameter of interest here) reflects the slope of the reaction function, with parameterizes the degrees of interdependence in efficiency (i.e., how much a hospital's efficiency is influenced by the average efficiency of immediate (and more distance) neighbors). If efficiencies autocorrelate spatially, this model will purge the coefficients in β from some bias. However, it can also be the case that some of the spatial auto-correlation results from auto-correlation in the determinants (characteristics in X). To control for this, the second variant of the SRM, called the Spatial Error Model (SEM) regresses the first (spatial) difference of the efficiency scores on the first (spatial) difference of the determinants:

$$(I_n - \gamma W)y = (I_n - \gamma W)X\beta + \varepsilon$$

We follow standard practice in the field and estimate both variants (SLM and SEM) and choose the best fit based on the Lagrange Multiplier (LM) test (Anselin 2013). Because the LM test led us to choose the SLM (statistic of 4.405, significant at the 5% level) over the SEM (statistic of 0.87, not significant at the 10% level), we detail below the estimation strategy followed for the SLM.

2.5 Model specification and estimation of the SLM

It is quite likely that $W\theta$ is not exogenous and an OLS estimator of the coefficients would be biased and inconsistent. To address the endogeneity bias, we instrument $W\theta$ by the first and second orders of spatial lags of explanatory variables, WX and W^2X, as suggested by Kelejian and Prucha (1998) and Kelejian and Robinson (1993). We then use the Robust Kleinbergen-Paap rk Wak *F* statistics to test for weak instruments and Sagan-Hansen *J* test for over-identification.

We then ran a spatial two-stage least squares (2SLS) using those instruments; Anselin (1999) shows that spatial 2SLS achieved the consistency and asymptotic normality properties of the standard 2SLS. We also ran a generalized method of moments (GMM) (Anselin 2006; Lesage & Pace 2009), as it is as efficient as a maximum likelihood estimation but requires much less computation time (Kelejian & Prucha 1998).

We compared the estimation results across various specifications (using binary and contiguity) and estimation methods (OLS, 2SLS and GMM) following Anselin and Bera (1998). We checked the sensitivity of our results to the definition of the catchment area by estimating the model using the distance threshold at 60km, 120km, 180km and 240km respectively.

2.6 Non-Spatial explanatory variables (X)

In this study, we want to estimate the effect of determinants of efficiency controlling for spatial auto-correlation. The list of potential determinants has been established by including the key factors in a previous study using the same data source (Wang et al. 2018). These determinants can be categorized into actionable and environmental factors. Actionable factors are: first, characteristics reflecting the ability of the hospital to cooperate with the rest of the health care system, such as the proportion of acute care transfers in total number of stays, the percentage of long term patients (to capture the effect of atypical cases), the percentage of patients discharged to long term care facilities, and the percentage of patients discharged to home care. A second sub-category of actionable factors is the hospital’s management strategy on resource allocation, represented by the number of nursing hours per weighted cases. We also control for size (number of beds, Asmild et al., 2013) and overall readmission rate as a proxy for quality of care (and quality of management).

Environmental factors are: the percentage of patients who live in a rural area, and whether the hospital is a teaching hospital since teaching hospitals have high unit costs associated with treating complex patients which may affect their efficiency (Hollingsworth 2003).

3. Results

3.1 Results of the efficiency scores

Table 1 shows descriptive statistics for all variables included in the empirical analysis. The average hospital spends $94.7 million on acute inpatient care, to treat 14,200 case-mixed adjusted cases, or $6,667 per case-mixed adjusted stay. The Hospital Standardized Mortality Ratio ranged from 62 to 183 across hospitals. There is also variation in the value of explanatory variables across hospitals: the average proportion of long-stay patients is 5% but some hospitals have 14% of their stays as long stay ones, 9% of patients are transferred but it goes as high as 21%; similarly, when the average hospital uses 45.6 nursing hours per weighted case, the largest number of hours is 116 hours. Some hospitals serve urban patients only and some rural patients only.

Table 1 insert here.

Table 2 shows some moments of the distribution of efficiency scores: Canadian hospitals reach an average efficiency level of 0.69: It indicates their output (quantity and quality) could be improved by 31%, without changing the inputs if they operated efficiently. It also shows that the average distance between two nearest hospitals in Canada is 62 km, some being in the same city and some being as far as 789 kilometers from their nearest neighbour.

Table 2 insert here.

The Moran's *I* test shows that the statistic is highly significant ($p<0.006$) suggesting strong spatial correlation. This is true for all values of *d*. It is clear that spatial independence should be taken into account at the second stage of analysis in order to obtain unbiased estimates of the determinants of efficiency.

Considering that the average distance of a hospital to its first nearest neighbor is 62km, as well as that the average number of hospitals neighbors (n=11) falling into the distance band [0, 120km], we take a value of d=120km as our baseline model. This restricted our sample to 156 hospitals (instead of 179), since 23 hospitals do not have any neighbors within 120km distance.

In Table 2, the weighted average efficiency of immediate hospital neighbors($w\theta$), on average, increases to 0.72 when using the contiguity weighting matrix with the distance threshold at 120 km (d<120km) but using provinces boundaries leaves average efficiency at 0.69 while reducing dispersion.

3.2 Spatial interdependence of the efficiency of acute inpatient care

Tables 3 and 4 present results from the SLM model with the contiguity and binary spatial weighting matrices. In Table 3, the first model is the simple OLS that ignores any potential interdependence (column 1). Columns 2 to 4 present the results of models with interdependence: a simple OLS, the SLG-2SLS and the SLG-GMM. All three models with interdependence show a significant and positive effect of spatial lag: if a hospital's neighbors are more efficient than the average by 10%, the hospital efficiency increases by approximately 3%, suggesting that hospitals are complements to each other in Canada. Tests for weak IV or over-identification confirm the quality of the IV estimates and the F-values for the Breusch-Pagan test are 18.30 for the 2SLS/GMM methods, indicating no heteroscedasticity in the data.

Table 3 insert here.

If using the spatial binary weighting (0<*d*<120km), the estimated spillovers are similar but the magnitude change slightly (Appendix A2). However, if the spatial binary weighting matrix changes using the province as the boundary, there is no significant spatial interdependence of hospital efficiency.

To test the sensitivity of the results to the parameter threshold, we estimate the SLM-GMM using three alternative neighborhood matrices based on *d*=60km, 180km and 240km. The corresponding sample sizes are 126, 171 and 172 hospitals which have at least one neighbor hospital within each distance. Table 4 indicates that the magnitude of the positive spillovers decreases when the value of d increase, from 0.357 (*d*<=60 km) to 0.211 (*d*<=180km).

Table 4 insert here.

3.3 Determinants of the efficiency of acute inpatient care

Table 3 and 4 also present the estimated coefficients of the explanatory variables on hospital efficiency controlling for spatial interdependence. Adding spatial interdependence improves the adjusted R^2, from 58% to 62% but does not change the estimated coefficients of the main determinants of efficiency.

Hospitals with more long term patients are less efficient: decreasing the proportion of long stay patients by 1% would increase the productive efficiency of acute inpatient care by approximately 2%. This confirms the results of Varabyova and Schroyogg (2013), who used OECD countries as data points to identify a production frontier of inpatient care and concluded that Canada was more efficient, with shorter average stays, than comparable countries. Transferring patients also improves efficiency, with an elasticity of 0.7. Larger hospitals (more beds) are more efficient, as they benefit from economies of scale, a result often found in the literature (Nwagbara et al. 2016; Preyra & Pink 2005), with an elasticity of 0.4. Other determinants have no significant impact on efficiency.

4. Conclusion and Discussion

This is the first study of the effect of spatial interdependence of hospital on the productive efficiency of acute inpatient care in Canada. Efficiency scores are estimated using a bootstrapped DEA and the effects of determinants of efficiency are estimated using a SLM to control for spatial interdependence. We find a strong and significant positive effect of spatial interdependence (elasticity of 0.3): Canadian hospitals are clearly complements to each other and work in networks much more than in competition. This finding is robust to choices made on the maximum distance for spillover but spatial interdependence becomes not significant when provincial boundaries set the limit to influence, (the geographically weighed matrix is based on the boundaries of a province). Provinces are responsible for planning and funding hospital healthcare services. Differences in sources of financing, payment mechanisms, supply of health services and benefits packages across the provinces may make collaboration difficult and competition weak.

This positive spillover effect of efficiency across hospitals may result from various institutional characteristics of the Canadian hospital sector.

Firstly, hospitals in Canada chose cooperative strategies rather than competition to improve their technical efficiency through transferring the patient on the basis of their need. Mostly likely it is because they are paid an annual global budget to serve a local population; the government provides the funding and expect results, both in quantity (enough throughput to avoid long wait times) and quality (no adverse events). Governments monitor wait lists and adverse events (and are held accountable for these by the public), and hospitals are better off cooperating to reduce both, mostly using transfers, rather than working against each other. Also, hospitals are subject to benchmarking and emulate each other and it is easy to imagine that hospitals tend to compare to their local peers. Last, hospitals in similar areas share resources that we cannot observe, such as human resources (human capital of hospital workers), practice styles of physicians, and information on local population health: rather than interdependence as a peer effect, spatial auto-correlation could result from unobservable heterogeneity at the area level.

We also find that transfers (in or out) have a positive association with efficiency, confirming that cooperation among hospitals improves efficiency.

From a policy perspective, our results suggest that the sharing of information on hospital outputs can improve overall efficiency of acute inpatient care, as would encouraging the transfer of information to transfer patients. This is true within the hospital sector but also between hospitals and long-term care facilities or rehabilitation facilities. Fragmentation is a source of inefficiency. This also raises fairness concerns, though, as Canada might well end up with large variations in efficiency across geographic clusters.

A limit of the study is that we do not have solid data on transfers, that would allow us to use the intensity of patient flows between hospitals instead of geographic distance to construct the spatial weighting matrix. In Canada, most transfers among hospitals are for routine, non–life-threatening reasons, using the Emergency Medical Services (EMS) system. The structure of emergency services varies greatly from province to province. For example, Alberta is centralizing its EMS structure to transfer responsibility to Alberta Health Services by April 2009. At the other end of the spectrum, emergency services in Nova Scotia are regulated by the Department of Health but managed by a private company, Emergency Medical Care. Therefore, it is difficult to make comparisons due to the diversity of governance structures and administration of EMS across Canada.

At the other side, it's well-established that any intra- or inter-hospital patient transfer should aim at maintaining optimal health of the patient which is carried out by transferring the patient to the nearest facility providing highest specialised care (Iwashyna and Courey, 2011). Robinson et al. (2009) pointed out that the majority of patient transfers in Ontario are for non-urgent reasons and for short distances. The results indicate that the distances between hospitals could be a valid proxy for the volumes of patient transfers. Using the distance weighing matrix could effectively and accurately detect the spatial interdependence among the hospitals with respect to the technical efficiency.

Funding Source:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References:

Anselin, L. 1988b. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Anselin, L. 1999. *Spatial econometrics, An updated review*. Regional Economics Applications Laboratory (REAL), University of Illinois, Urbana-Champaign.

Anselin, L 2006 "Spatial econometrics", in Mills, TC & Patterson, K (eds), *Palgrave Handbook of Econometrics: Econometric Theory*, Palgrave Macmillan, Basingstoke, pp. 901–969.

Anselin, L 2013, *Spatial econometrics: methods and models,* 4th edn, Springer Science & Business Media, Amsterdam.

Anselin, L., and A. Bera 1998, "Spatial dependence in linear regression models with an introduction to spatial econometrics", In A. Ullah and D.E.A. Giles (eds.) *Handbook of Applied Economic Statistics*, pp. 237–89. New York: Marcel Dekker

Asmild, M, Hollingsworth, B, & Birch, S 2013. "The scale efficiency of hospitals: One size does not fit all." *Journal of Productivity Analysis*, vol. 40, no. 2, pp. 197-206. DOI: 10.1007/s11123-012-0332-9

Canadian Institute for Health Information (CIHI). 2013. *MIS Standards 2013*. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2016a. *Indicator Library: Peer Group Methodology*. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2016b. *Hospital Standardized Mortality Ratio: Technical Notes*. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2017. *DAD Resource Intensity Weights and Expected length of Stay*. Ottawa, ON: CIHI.

Cliff, A & Ord, JK 1973, *Spatial Autocorrelation*, Pion Ltd, London

Cliff, A. and Ord, J. K. 1981. *Spatial Processes: Models and Applications*. Pion Ltd, London.

Cooper, WW, Seiford, LM & Zhu, J.. 2011. *Handbook on Data Envelopment Analysis: International Series in Operations Research& Management Science*, 2nd edn, Springer, New York.

Cooper Z, Gibbons, S, Jones, S, & McGuire, A 2011. "Does hospital competition save lives? Evidence from the English NHS patient choice reforms." *The Economics Journal*, vol. 121, no. 554, pp. F228- F260. DOI: 10.1111/j.1468-0297.2011.02449.x

Currie, J, MacLeod, WB, & Van Parys, J.. 2016. "Provider practice style and patient health outcomes: the case of heart attacks." *Journal of Health Economics*, vol. 47, pp. 64-80. DOI: 10.1016/j.jhealeco.2016.01.013

Dormann, Carsten F., Jana M. McPherson, Miguel B. Araújo, Roger Bivand, Janine Bolliger, Gudrun Carl, Richard G. Davies et al. 2007, "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review." *Ecography* 30, no. 5, pp. 609-628.

Gaynor, M, Propper, C, & Seiler, S. 2016. "Free to choose? Reform, choice, and consideration sets in the English National Health Service." *American Economic Review* vol. 106, no. 11, pp. 3521-3557. DOI: 10.1257/aer.20121532

Gravelle, H, Santos, R, Siciliani, L 2014. "Does a hospital's quality depend on the quality of other hospitals? A spatial econometrics approach." *Regional Science and Urban Economics* vol. 49, pp. 203–216. DOI: 10.1016/j.regsciurbeco.2014.09.005

Getis, A, Mur, J, Zoller, HG (eds.) 2004, *Spatial Econometrics and Spatial Statistics*, Palgrave MacMillan, New York.

Herwartz, H, & Strumann, C 2010. "Spatial Interdependence of Hospital Efficiency in Germany." paper presented to Beitrage zur Jahrestagugng des Vereins Fur Socialpolitik 2010: Okonomie der Familie, Session: Public Policy Spillovers, No. F7-V2

Ferrier, GD & Valdmanis, VG 2006. "Peer effects and efficiency: the influence of competitors' performance on hospital efficiency." paper presented to Economics of Population Health: Inaugural Conference of the American Society of Health Economists, Madison, WI, USA, Jun 04, 2006.

Hollingsworth, B. 2003. "Non-parametric and parametric applications measuring efficiency in health care." *Health care management science*, vol. 6, no. 4, *pp.* 203-218. DOI: 10.1023/A:1026255523228

Hollingsworth, B. 2008. "The measurement of efficiency and productivity of health care delivery." *Health Economics*, 17: 1107-1128

Iwashyna TJ & Courey AJ 2001. "Guided transfer of critically ill patients: Where patients are transferred can be an informed choice." *Current Opinions in Critical Care,* vol. 16, no. 6, pp. 641-647. DOI: 10.1097/mcc.0b013e32834b3e55

Kelejian, HH, & Robinson, DP 1993. "A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model." *Regional Science* vol. 72, no. 3, pp. 297–312. DOI: 10.1111/j.1435-5597.1993.tb01878.x

Kelejian, HH, & Prucha, IR 1998. "A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances." *Journal of Real*

Estate Finance and Economics vol. 17, no.1, pp. 99–121. DOI: 10.1023/A:1007707430416

LeSage, J, & Pace, RK 2009. *Introduction to Spatial Econometrics*, CRC Press, Boca Raton.

Lisi, D, Moscone, F, Tosetti, E, & Vinciotti, V 2017. *Hospital interdependence in a competitive institutional environment: Evidence from Italy* (No. 17/07). HEDG, c/o Department of Economics, University of York.

Longo, F, Siciliani, L, Gravelle, H, & Santos, R 2017. "Do Hospitals Respond To Neighbours' Quality and Efficiency? A Spatial Econometrics Approach." *Health Economics*, Forthcoming.

Mobley, LR 2003. "Estimating hospital market pricing: an equilibrium approach using spatial econometrics." *Regional Science and Urban Economics* vol. 33, no. 4, pp. 489-516. DOI: 10.1016/s0166-0462(02)00059-5

Mobley, LR, Frech, T, & Anselin, L 2008. "Spatial interaction, spatial multipliers, and hospital competition. Economics." Working Paper Series 07-08, Department of Economics, UC Santa Barbara.

Moran P. 1948. "The interpretation of statistical maps." *Biometrika* 35, 255-260.

Moran P.1950a. "Notes on continuous stochastic phenomena." *Biometrika* 37, 17-23.

Moran P. 1950b. "A test for the serial dependence of residuals." *Biometrika* 37, 178-181.

Norton, EC, Van Houtven, CH, Lindrooth, RC, Normand, SLT, & Dickey, B 2002. "Does prospective payment reduce inpatient length of stay?" *Health Economics* vol. 11, no. 5, pp. 377-387, DOI: 10.1002/hec.675.

OECD. 2011. "OECD Guidance, Competition Assessment Toolkit." *OECD* Vol 2, Version 2.0.

Pai, CW, Ozcan, YA, & Jiang, HJ 2000. "Regional Variation in Physician Practice pattern: an examination of technical and cost efficiency for treating sinusitis." *Journal of Medical Systems* vol. 24, no. 2, pp. 103-117, DOI: 10.1023/A:1005568814163.

Preyra, C & Pink, G 2005. "Scale and scope efficiencies through hospital consolidations." *Journal of Health Economics*, vol. 25, no. 6, pp. 1049-1068, DOI:1016/j.jhealeco.2005.12.006

Propper, Carol. 2012. "Competition, Incentives and the English NHS." *Health Economics*, 21(1).

Redding, SJ & Rossi-Hansberg, E 2017. "Quantitative Spatial Economics." *Annual Review of Economics*, vol. 9, pp. 21-58. DOI: 10.1146/annurev-economics-063016-103713

Robinson, V, Goel, V, MacDonald, RD, & Manuel, D, 2009. “Inter-Facility Patient Transfers in Ontario: Do You Know What Your Local Ambulance Is Being Used For?” *Healthcare Policy*, vol. 4, no. 3, pp. 53–66. DOI: 10.12927/hcpol.2009.20478

Simar, L., and P.W. Wilson. 1998. “Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models.” *Management Science* 44(1):49-61.

Simar, L., and P.W. Wilson. 2007. “Estimation and inference in two-stage, semi-parametric models of production processes.” *Journal of Econometrics* 136(1): 31-64.

Shahid, R, Bertazzon, S, Knudtson, ML, & Ghali, WA 2009. “Comparison of distance measures in spatial analytical modeling for health service planning.” *BMC Health Services Research*, vol. 9, no. 1, pp. 1-14, DOI: 10.1186/1472-6963-9-200

Sheikhzadeh, Y, Roudsari, AV, Vahidi, RG, Emrouznejad, A, & Dastgiri, S. 2012. “Public and private hospital services reform using data envelopment analysis to measure technical, scale, allocative, and cost efficiencies.” *Health Promotion Perspectives*, vol. *2*, no. 1, pp. 28. DOI: 10.5681/hpp.2012.004

Srinivasan, S 2008. Sptial Regression Models, in Shekhar, S, Xiong, H, Zhou, X (eds.) *Encyclopedia of GIS*, Springer, pp. 1102-1105

Tobler W., 1970. "A computer movie simulating urban growth in the Detroit region." *Economic Geography*, 46(Supplement): 234-240.

Varabyova, Y & Schreyögg, J 2013. "International comparisons of the technical efficiency of the hospital sector: panel data analysis of OECD countries using parametric and non-parametric approaches." *Health policy* vol. 112, no. 1, pp. 70-79. DOI: 10.1016/j.healthpol.2013.03.003

Wang L., Grignon M. Perry S., Chen X., Ytsma A., Allin S., and Gapanenko K. 2018. “The Determinants of the Technical Efficiency of Acute Inpatient Care in Canada.” *Health Service Research*, 53(6):4829-4847. doi: 10.1111/1475-6773.12861.

Wilson, P. W. 2008. “FEAR 1.0: A Software Package for Frontier Efficiency Analysis with R.” *Socio-Economic Planning Sciences* 42, 247--254.

Table 1: Descriptive statistics of the variables used in the analysis

			Range	
Input:	Mean	Std.	Min.	Max.
Inpatient Cost ex. mental and rehabilitation care (100,000')	947	1159	83	6410
Output:				
Case mixed RIW (100')	142	169	10.83	817.47
Hospital Standardized Mortality Ratio	104	19	62	183
Explanatory variables				
% Long stay patients	5.1	2.5	1.45	14.00
% Acute patients transfers (in or out)	9.1	4.3	2.36	21.46
Beds (log form)	5.2	0.9	3.40	7.22
Nursing hours per weighted case	45.6	11.9	32.23	116.26
% Rural patients	26.2	21.7	0.07	97.80
% Transferred to long term care for first	2	1.3	0.00	20.30
% Overall readmission rate	8.8	0.9	5.74	11.20
% Discharged home with supportive	10.7	7.1	0.00	31.14

Table 2: Summary statistics of the distance and the spatially lagged efficiency score (N=179)

Variable	Mean	Std.	Min.	max.
Distance from each DMU to its first nearest DMU	62	99	0.3	789
Number of neighbors falling in the distance band: $0<d_{ij}<=120$km	11.21	11.81	0	37
Robust efficiency score (θ)	0.697	0.119	0.338	0.957
$W\theta$ (Binary weighted matrix if belong to same province)	0.697	0.059	0.478	0.746
$W\theta$ efficiency score(Contiguity weighted matrix if d<120km)	0.716	0.093	0.37	0.89

Note: The prefix W indicates spatial lag of the corresponding variables.

Table 3: Estimating the determinants of efficiency, controlling for spatial interdependence. Three main models (OLS, 2SLS and GMM) compared to OLS without spatial interdependence. Parameter *d* set at 120km using contiguity weighting matrices.

	OLS	**OLS with interdependence**	**IV 2SLS**	**IV GMM**
	Coef./t	**Coef./t**	**Coef./t**	**Coef./t**
W efficiency score		0.317***	0.297**	0.303**
Percent Long Stay Cases	-0.020***	-0.015***	-0.015***	-0.015***
Beds (Log)	0.039***	0.040***	0.035***	0.036***
Average nursing hr per weighted case	-0.002***	-0.003***	-0.003***	-0.004***
Percent acute Transfer	0.007***	0.009***	0.008***	0.009***
Percent home care transfer	-0.004	-0.003	-0.003	-0.003
Percent long term care transfer first time	0.005	0.003	0.003	0.003
Percent of rural patients	-0.001	-0.000	-0.000	-0.001
overall readmission rate	0.008	0.004	0.004	0.002
Teaching status	0.033	0.031	0.031	0.038
Constant	0.572***	0.375***	0.425**	0.455***
Sample size	179	156	156	1560
R^2	0.575	0.629	0.629	0.624
Weak IV			11.1974	11.1974
Overidentification (*p*-value)			18.2925	18.2925

Note: The prefix W indicates spatial lag of the corresponding variable. The instruments for spatial lag of dependent variable is ***WX***, ***W*2*X***. Weak IV is the robust Kleinbergen–Paap *rk* Wald *F* statistic for test of weak instruments. Overidentification shows *p*-value of Sagan–Hansen *J* test; ***p<0.01, **p<0.05

Table 4: SLM-GMM model with different distance threshold

	d<=60km	*d*<=120km	*d*<=180km	*d*<=240km
Contiguity weighted matrix				
W efficiency score	0.357***	0.303***	0.211**	0.215**
# of hospitals without neighbors	53	23	8	7
R^2	0.648	0.624	0.616	0.589
Binary weighted matrix				
W efficiency score	0.304***	0.323**	0.238**	0.249**
R^2	0.651	0.626	0.613	0.581

Note: The prefix W indicates spatial lag of the corresponding variable; The other explanatory variables are as same as the models in Table 3; *** p<0.01, ** p<0.05

Appendix:

A1: Equation Used to Calculate the Distances between Two Hospitals Using the Latitude and Longitude of Two Hospitals.

$$D = 6{,}370{,}997 * \arccos[\sin(Lat_A) * \sin(Lat_B) + \cos(Lat_B) * \cos(Lat_B) * \cos(Long_A - Long_B)]$$

where

- D = distance (in metres)
- Lat_A = latitude of point A (in radians)
- $Long_A$ = longitude of point A (in radians)
- Lat_B = latitude of point B (in radians)
- $Long_B$ = longitude of point B (in radians)
- arcos = arc cosine
- cos = cosine
- sin = sine
- 6,370,997 is the radius of the sphere (in metres)

A2: Estimating the determinants of efficiency, controlling for spatial interdependence. Three main models (OLS, 2SLS and GMM) compared to OLS without spatial interdependence. Parameter d set at 120km with binary weights matrices.

	OLS	**OLS with interdependence**	**IV 2SLS**	**IV GMM**
	Coef./t	**Coef./t**	**Coef./t**	**Coef./t**
W efficiency score		0.325***	0.267***	0.322***
Percent Long Stay Cases	-0.024***	-0.016***	-0.017***	-0.015***
Beds (Log)	0.032**	0.027**	0.028**	0.024**
Average nursing hour per weighted case	-0.002***	-0.003***	-0.003***	-0.004***
Percent acute Transfer	0.004**	0.005**	0.005**	0.005**
Percent home care transfer	-0.002	-0.001	-0.001	-0.001
Percent long term care transfer first time	0.006	0.001	0.002	0.003
Percent of rural patients	0.000	0.000	0.000	0.000
overall readmission rate	0.001	0.003	0.002	0.006
Teaching status	0.033	0.033	0.033	0.047
Constant	0.773***	0.461***	0.514***	0.455***
$\mathbf{R^2}$	57.4	63.0	62.9	62.6

Note: The prefix W indicates spatial lag of the corresponding variable. The instruments for spatial lag of dependent variable is $\boldsymbol{WX}$, $\boldsymbol{W}2\boldsymbol{X}$.

CHAPTER IV: THE TECHNICAL AND SCALE EFFICIENCY OF SMALL HOSPITALS IN CANADA

ABSTRACT

Background: Small hospitals play an important role in the provision of health care for the population residing in rural or remote areas in Canada (approximately 20% of the 35 million Canadians). This role and, as a result, the production function of small hospital differs from that of larger hospitals situated in urban centres, and the efficiency of small hospitals cannot be assessed with the same criteria. For instance, whereas size is one determinant of efficiency among large hospitals, it is a crucial one for small hospitals, and requires special treatment. This paper explores the technical and scale efficiencies of small hospitals, as well as the factors that influence them, for Canada in 2012/2013.

Methods: We use cross-sectional data from the Canadian Management Information System (MIS) database (CMDB) and Discharge Abstract Database (DAD). We estimate output-oriented technical efficiency with Variable Returns to Scale (VRS), as well as scale efficiency using a Data Envelopment Analysis (DEA) with FEAR package in Rstudio. Using a double bootstrap Tobit model DEA, efficiency scores are regressed against selected institutional and contextual/environmental factors to estimate their impacts on technical and scale efficiency respectively.

Results: On average, the average VRS productive efficiency score is 54 % and the scale efficiency score is 66 %, meaning that small hospitals are far from their frontier either in production or scale. Technically inefficient hospitals could become more efficient by improving the ability to discharge patients to home care or collaborate with the neighbor urban hospitals through more patient transfers. The contextual effects, such as the average income and the obesity rate in the region where the small hospitals were located, are significantly related to the technical efficiency. The small hospitals with higher cost shares of diagnostic costs or nursing were more scale efficient.

Conclusions: The small hospitals identified at the high and low extremes of technical or scale efficiency should be investigated further to determine how and why production processes are operating differently at these hospitals. The policy maker would aim to enforce the collaboration with the urban hospitals and the post-acute home care, as well as to develop context-appropriate strategies for supporting hospitals with low technical efficiency to improve their service and thereby better address unmet needs for hospital services in Canada.

Keywords: Small hospital, Data envelopment analysis, Scale efficiency, Technical efficiency, Tobit model, Canada

Introduction

Background: Small hospitals in Canada

A small hospital is defined as a hospital delivering less than a certain number of weighted inpatient cases per year, with weights reflecting intensity of resources per stay (Resource Intensity Weights, established by the Canadian Institute for Health Information (CIHI, 2017)). The threshold ("certain number") varies across definitions, with CIHI setting it at 2,000 (this is approximately 5 stays per day, CIHI, 2016) and the Ontario Ministry of Health and Long-Term Care at 4,000 (or 10 stays per day). In this study, we use the CIHI threshold and focus on very small hospitals, delivering less than 5 cases in a given day.

In Canada, as elsewhere, the major part of acute inpatient care services is delivered by non-small (teaching, large, and medium) hospitals. However, Canada relies on a substantial number of small hospitals (332 in the country, not counting Quebec and Nunavut for which information is not available or comparable, out of a total of 722 hospitals of all sizes) to deliver care. Most of these small hospitals are located in areas that can be considered rural, based on the definition of rural by Statistics Canada, i.e., an area with less than 1,000 inhabitants and population density lower than 400 per square kilometer (Bollman and Clemenson, 2008). As a result, the terms "small hospital" and "rural hospital" are often used interchangeably. According to the 2011 Canadian census, around 20% of Canadians lived in such rural areas. A specificity of Canadian geography is the existence of so-called remote rural areas, or large and sparsely populated areas with no major urban centre. Distance, more than rurality per se, may be the real reason why Canada has so many small hospitals, as they serve as the foundations of a rural health care delivery system that allows rural and remote residents to access acute inpatient care (Rechel et al., 2016): 22.5% of Canadians live more than one hour driving distance from a level I or II trauma centre (Fleet et al., 2013) and 17% of rural women are more than two hours driving distance from the hospital of birth (CIHI, 2013).

What challenges do small hospitals face?

Due to the relatively low level of demand for their services, small/rural hospitals tend to have few beds, low occupancy rates (relative to the high occupancy rates of hospitals in Canada in general); on the supply side, they generally rely on fewer skilled nurses (or skills in general) and less technology to deal with complex cases compared to larger/urban hospitals. In order to meet the need of health care of rural residents, small hospitals need to play various roles, including in-patient beds for medicine and surgery, a full service 24/7 Emergency Department, a variety of basic clinics and a full service laboratory and diagnostic imaging departments. Small hospitals often act essentially as community health centres that triage acute patients (and transfer them to larger hospitals) and provide post-acute recovery care to patient discharged from larger hospitals. Faced with such constraints, small hospitals cannot be expected to perform at the same level of efficiency or complexity as urban hospitals. Small/rural hospitals struggle with the challenges of delivering efficient and effective high quality health services in an increasingly technological and complex health system. Also, a large component of the services they provide pertain to nursing home care, meaning a large proportion of long-stay for elderly patients causing large operating costs.

One specific challenge is directly linked to size and the inability to benefit from economies of scale that would allow them to bring their average cost down. Another separate issue is that the cost of labour and supplies can be considerably higher in rural and remote areas. In order to protect these hospitals that serve unique needs in rural communities and also because, in most instances, the local hospital is the main employer in the area, governments choose to over-compensate small hospitals for the numbers of cases they treat. For example, Ontario excludes its small hospitals from any attempt at introducing activity-based funding in the funding formula and keeps them entirely under global budget based on historic funding figures, regardless of the number of patients they treat and the quality of care they provide (MOHLTC, 2017). British Columbia provides additional funds to small and remote hospitals relative to other hospitals. Other countries with large numbers of small hospitals, such as Australia and the United States, also acknowledge the fact that small and remote hospitals cannot operate as large hospitals and fund them through global budgets or cost-plus reimbursement policies (Sutherland et al., 2013, Rechel et al., 2016). It is therefore crucial for these jurisdictions to better understand the trade-off between efficiency and equity they face when they support these small hospitals, as well as to understand how the deadweight loss can be minimized.

What types of efficiency are of most concern to policy makers?

Two main questions need to be addressed in order to make sure that supporting small hospitals is not too costly in scarce resources needed elsewhere in the health care system. The basic question is that of productive efficiency, which refers to the ability of a small hospital to transform inputs into outputs; productive inefficiency is the distance of a given hospital to the production frontier. It is sometimes addressed as the question of cost-efficiency, or how far is the hospital from the cost-frontier. By identifying the most efficient small hospitals and learning from them, policy-makers could devise interventions that would bring more outputs for the same level of input (or spending). The second question is that scale efficiency: how small can a small hospital be, or is there an optimal size within small hospitals (not too small, but not too large either given they operate in rural and remote areas), based on their input-output mix? Determining the optimal size of a small hospital would help policy makers make decisions about downsizing, mergers or hospitals closing. Moreover, scale efficiency is an important determinant of overall efficiency and knowing the latter will help decision-makers understand better the determinants of productive efficiency, conditional on scale. The issue of scale efficiency is particularly striking among small hospitals because their average occupancy rate is lower than urban/larger hospitals and the demand for their services is more variable over time.

What do we know about technical and scale efficiency within small hospitals?

Whereas research on efficiency among urban hospitals has become widespread, the study of small/rural hospitals is still in its infancy, in particular in Canada. Most studies are content with pooling all hospitals together, independent of size (small, large, teaching and medium size urban hospitals), and using size as a determinant of efficiency (often by comparing efficiency by categories of size) (Asmild et al., 2013). Canadian findings on size and efficiency vary across studies and, possibly, contexts: Two studies of Ontario hospitals separated by ten years (Gruca & Nath, 2001 and Chowdhury et al., 2011) find either that urban hospitals are less (Gruca & Nath, 2001) or more (Chowdury et al., 2011) efficient than small and rural ones. The former result is an artefact of the specification used, in which long-term care days contribute to the output of hospitals (and small hospitals produce much more long-term care days than urban ones). A study focused on small and rural hospitals is warranted, to provide a better and more meaningful analysis of the determinants of efficiency within that category, including the role of size among small hospitals. The Manitoba Center for Health Policy and Evaluation conducted such a study of the performance of rural and northern hospitals (Stewart et al., 2000) and found hospitals performing well on a variety of indicators might be used as benchmarks to target. Their findings identified occupancy rate, use relative to need ratio, and the intensity of services as determinants of efficient practices.

What are the objective/questions for this study?

The overall objective of this study is to evaluate the level of technical and scale efficiencies of small hospitals across Canada. To achieve this objective, we address three questions: (1) What was the average level of (technical) efficiency of small hospitals in Canada in 2012-2013, in terms of case mix weighted inpatient care and day surgery, as well as the combined emergency department and clinical visits? (2) How small should small hospitals be to achieve their optimal size based on their current output? (3) How do technical and scale efficiency scores correlate with the institutional and contextual factors of a small hospital?

What is the added value of this study?

This study explores novel ways of using administrative data to evaluate both the technical and scale efficiency of small hospitals across Canada: The Canadian Institute for Health Information (CIHI) provides rich data on individual hospitals in Canada, including financial expenses and operation variables (such as the number of discharges and the outpatient visits) and we link these administrative data to contextual variables, such as the average income, and the prevalence of obesity in the population of the catchment area of the hospital, derived from the self-report 2006 census and the 2012 Canadian Community of Health Survey(CCHS). This allows us to better understand the determinants of efficiency and suggest avenues for improvement that can be taken up by policy makers to develop benchmarks and establish a standard to get better outcomes from small and rural hospitals. Second, the scope of this research is not restricted to one province or small jurisdiction but expands across Canada (except the province of Quebec and the territories), allowing us to compare efficiency across various political contexts (but operating under a common set of values, the Canada Health Act).

Method and DATA

Technical and scale Efficiency concepts

We use an output orientation for our analysis: what is the level of output a DMU (here, a small and rural hospital) produce, relative to the maximum quantity of outputs it could produce with the same level of inputs? Output-orientation is consistent with most studies of efficiency in universal health care systems (e.g., for Canada, Wang et al., 2018, Allin et al., 2016), where DMUs are given a fixed quantity of resources (inputs) and asked to produce as much output as possible. An input-oriented measure of efficiency would assume Canadians receive exactly what they need from small hospitals and would seek for ways to cut their budgets without affecting services.

Technical efficiency can be broken down into pure technical efficiency and scale efficiency. Scale efficiency is a measure of the extent to which a hospital deviates from its optimal scale (defined as the portion of the input-output curve in which there are constant returns to scale in the relationship between outputs and inputs). Scale efficiency is of concern to decision-makers because, in the Canadian context, the budget of each hospital is determined by the Ministry of Health and Long-Term Care (MOHLT), therefore the size of each hospital (amount of inputs, mostly labour) is set by decision-makers, rather than the hospital itself. Pure technical efficiency is the portion of efficiency that cannot be attributed to deviations from optimal scale (scale efficiency) and is therefore estimated as a residual (total efficiency minus scale efficiency).

We follow a non-parametric approach, called Data Envelopment Analysis. DEA determines efficient acute hospitals by comparing the actual level of outputs of one hospital with the simulated level of linear combinations of other hospitals in the data set. More specifically, one hospital is inefficient to some extent if a linear combination of two other hospitals observed in the data that would use the same level of inputs can produce more. The level of inefficiency is the difference between actual output and this simulated output of a linear combination of other hospitals that would have the same level of input. There are two types of DEA models. One, developed by Charnes, Cooper and Rhodes (CCR) (1981) and the other, proposed by Banker, Charnes and Cooper (BCC) (1984). CCR assumes that production has constant returns to scale (CRS) indicating any change in the input (at any level of input) will result in a proportionate change in the output. BCC assumes that production has variable returns to scale (VRS) and therefore, the effect of an increase in input on output will vary with the initial level of input, decreasing when input increasing in the case of decreasing returns to scale, and increasing otherwise. Whether returns are decreasing or increasing along the whole range of inputs is determined during the procedure of linear combinations. Since BCC imposes fewer restrictions on the data, and CCR relies on an untestable assumption of CRS (on slope applies to all DMUs), we follow the VRS approach of BCC (Allin et al., 2015). A hospital is considered to be technically efficient if it scores one, whereas a score smaller than 1 implies that it is relatively technically inefficient, compared to peers in its efficiency reference set.

If a hospital operates at the optimal scale, the technical efficiency is equal to the scale efficiency, otherwise, the technical efficiency measured by the CCR model would be revised by the scale efficiency. The BCC model using VRS can incorporate the impact of scale efficiency in the evaluation of the technical efficiency. Thus, the scale efficiency can be calculated by dividing the CRS technical efficiency to VRS technical efficiency (Fare, Grosskopf and Lovell, 1994, pp. 73-74). Scale efficiency provides information on the optimal size of a hospital based on their input-output mix. If a hospital is scale inefficient, reducing or increasing its size (inputs) would result in more outputs per inputs. A scale efficiency score of one implies that the hospital is operating at optimal size. If the scale efficiency score is less than one, then the hospital is either too small or too big relative to its optimal size.

Data Description

Our empirical results are based on all small hospitals providing acute inpatient care in Canada outside of Quebec and the territory of Nunavut, for fiscal year 2012-2013 (April 2012 to March 2013). Inpatient episodes of care are extracted from the Acute Care Discharge Abstract Database (DAD), a standardized set of information provided by provinces and harmonized and made accessible by CIHI. It is comprised of records for each inpatient stay (at the time of discharge), with information standardized across provinces on age, sex, diagnoses, procedures received and length of stay (dates of admission and discharge of the patient), as well as transfers (to and from hospitals or other institutions). The quantity of outpatient care, including day procedures, emergency care and clinical visits is extracted from the National Ambulatory Care Reporting System (NACRS). NACRS contains data for all hospital-based and community-based ambulatory care, including day surgery, outpatient and community-based clinics, and emergency departments, and we select contacts pertaining to small and rural hospitals only.

The other data source is the financial data of the Canadian Management Information System (MIS) database (CMDB) for the fiscal year 2012-2013. CMDB consists of financial and statistical information on Canadian hospitals, updated annually from data submitted by provincial and territorial ministries. CMDB collects information on hospital resources including nursing, administration, diagnostic, therapeutic hours and research and education (CIHI, 2013). In our study, we use total budget instead of the detailed information on quantity of resources, again restricted to small hospitals. Merging the two sets of information together on the basis of a common hospital identifier, outputs can be related to inputs (budget) on a hospital basis.

DMUs: Small Hospitals

Based on CIHI's definition of small hospitals (less than 2,000 weighted inpatient cases), 332 small hospitals have been selected from DAD and CMDB. First, 50 small hospitals located in the urban area were excluded because they do not face the common challenges that small hospitals face, e.g., high level of transfers. Second, 13 small hospitals without any acute beds were dropped, since they are not really what can be called a hospital. Last, 30 small hospitals that strongly specialize in one specific type of care were excluded, because they do not offer the scope of treatments that most hospitals, especially small and rural hospitals, provide, and, as a result, cannot be compared meaningfully with the rest of the population of hospitals. These hospitals were identified using a Herfindahl–Hirschman Index (HHI) applied to the distribution of treatments by categories of service, provided by each hospital (OECD, 2011): if the HHI is greater than 0.25, the hospital is deemed specialized. In 20 out of these specialized 30 small hospitals, the majority of patients were hospitalized for rehabilitation, palliative care or awaiting placement, and for the other excluded 10 hospitals, specialization included pregnancy, delivery and newborn care, mental health, eye and respiratory system. Eventually, 229 small hospitals are included in the efficiency analysis.

Outputs

It is assumed that small and local hospitals produce four different types of output: number of inpatient stays, number of cases of day surgery, number of Emergency Department (ED) visits, and inverse of length of stay before transfer (if any) as a quality indicator. The first three outputs are quantitative, and the last one is qualitative, capturing the quality of stays provided by small hospitals. The indicators of quality of care used in the literature on efficiency have been developed primarily with urban hospitals in mind, such as all causes of readmission rate in 30 days, in-hospital sepsis rate, hospital standardized mortality ratio (HSMR, Wang et al., 2018) and length of stay (Brasel et al., 2007; Guru et al., 2005). Small hospitals and their environment differ from large and teaching urban hospitals, requiring different indicators of quality of care. For example, HSMR is meaningless among small hospitals because they usually provide very basic hospital services and deal with non-complex patients who are not likely to die. All causes for readmission in 30 days would signal poor quality of care for urban hospitals but is not suitable for small hospitals because one of core services is to triage and transfer complex cases to ensure that patients in rural and remote area to get access to care when needed (Rechel et al., 2017).

Baernholdt et al. (2010) suggest that the measure of quality of care needs to capture the unique features of rural/small hospitals, including how quickly a small hospital responds to a patient's need and the culture of the community. We thus use the inverse of the average of length of stay as a quality indicator in which the large value indicates a quick turnover of the acute bed as well as a short transfer time to another facility: we expect small hospitals to make quick decisions on where to transfers patients, if they cannot be discharged quickly.

Number of stays and day surgery cases are weighed by their resource intensity weight (RIW), a series of weights developed by CIHI, where each patient receives an RIW by the Case Mixed Group (CMG+) methodology that groups patients by diagnoses, and relative resource use and adjusts for age, sex, co-morbidities and interventions (CIHI, 2017). RIWs can be further adjusted for special cases (long stay outliers, transfers, death in the hospital or sign-out cases) called "atypical" weights (CIHI, 2013). However, in our analysis we use "typical" RIW which excludes the atypical cases because these adjustments provide resource for potentially inefficient care. We also excluded the weighted RIWs on the mental and rehabilitation care because they are not reported and weighted in a consistent way across provinces. The number of ED visits and clinical visits are not case-mixed adjusted. It must be noted that no indicator of quality is considered for the outputs of day surgery and outpatient visits. Jarvis (2016) proposed time from arrival to the triage for the ED visits as the quality indicator of the outpatient care, but the variable is not available in hospital databases in Canada and we could not use it as a result.

Inputs

Total operating expenditure is used as the measure of input; it aggregates various labour skill levels and all other types of resources, weighted by their unit costs. We exclude capital costs because they cannot easily be converted into yearly flows of expenditures. This latter exclusion is not too concerning because capital costs are small compared to operating expenses (CIHI, 2016). Sensitivity analyses including capital cost in the input variable were conducted to test that this methodological choice is not what drives the results. Consistent with the definition of output of inpatient care, cost for inpatient mental and rehabilitation care are also excluded from the measure of input.

Explaining efficiency through regression analysis

In the second stage, the efficiency scores (technical and scale) computed by the DEA are regressed against potential factors that can explain hospital efficiency, some, but not all, being under the control of the hospital. The choice of variables used as potential determinants in this paper has been guided by health care studies and data availability. The overall efficiency of the hospital is related to particular aspects of the service/case mix of the hospitals, their cost structure and human resources (Chilingerian and Sherman 1990, 2011). Specifically, the second stage used the following potential factors.

The percentage of long-stays among all stays, expected to be negatively related to efficiency (Xenos et al., 2017) since hospital resources remain committed to the same patient, to the detriment of new admissions.

The percentage of transfers to home care or long-term care facilities, which has been found to positively affect efficiency among large and urban hospitals, as they can discharge patients earlier than other hospitals (Wang et al., 2018).

The rate of transfers, since transferring patients to other acute facilities where they can receive the care they need is one of the key roles of a small hospital, and more transfers to or from acute hospitals are expected to lead to higher efficient (Gonzalez et al., 2013, Southard et al., 2005).

The percentage administrative costs, as it is a determinant of efficiency often cited in the literature (Sultan et al., 2018, Kalhor et al., 2016).

The spending on the diagnostic tests and nursing as the percentages of total hospital spending could affect efficiency because they reflect hospital management strategy on resource allocation (Mujasi et al., 2016, Chowdhury and Zelenyuk, 2016).

Demographic and socioeconomic characteristics of the population of the area where the hospital is located, including the percentage of residents with higher education and the prevalence of the obesity, as they have been shown to be strong determinants of hospital efficiency (Matranga et al., 2014, Mujasi et al., 2016, Fragkiadakis et al., 2016).

We estimated a Tobit model because DEA efficiency scores θ_i are bounded between 0 and 1. Due to serial correlation among the efficiency scores (relative) and the correlation of the inputs and outputs used in the first stage with second-stage environmental variables, we program the double bootstrap following Algorithm #2 in Simar and Wilson (2007) to estimate the association between the potential factors and the technical and scale efficiency of small hospitals respectively.

Double Bootstrapping Procedure

A main drawback of DEA is that it has no accommodation for noise or random error, as it uses a non-statistical approach (linear programming) to calculate the efficiency scores. Hence DEA is not able to provide a statistical foundation for the estimated efficiency frontier. In addition, the efficiency scores derived from the DEA are expected to be correlated with each other as the calculation of efficiency of one hospital incorporates all other hospitals in the same dataset. Thus, the assumption of independence of the random error term in the regression analysis is violated. To address this problem, a double bootstrapping procedure, proposed by Simar and Wilson (1998, 2007), was used in the second-stage analysis of this study, in which the bias-corrected estimate $\theta_i^{\hat{b}c}$ (bootstrap estimator) replace the θ_i and the bootstrap estimators were substituted for the crude estimators in the second-stage to calculate the standard error of the estimates. The detail bootstrap algorithm is presented in Appendix 1.

Model Specification

For each technical and scale efficiency, we estimated two models. The first is for the acute inpatient care:

Inputs: the cost on the inpatient care excluding capital, mental and rehabilitation inpatient care; Outputs: inpatient RIWs and the reverse of the length of stay.

The other one is for all types of services:

Input: total cost excluding capita, mental and rehabilitation inpatient care; Output: combined inpatient RIWs and day surgeon RIWs, combined ED visits and clinical visits, as well as the reverse of the length of stay.

RESULTS

The distribution of small hospitals across provinces

Table 1 presents the distribution of small hospitals by province, showing, not surprisingly, that most of small and rural hospitals are located in the least densely populated provinces known as the prairies (Alberta, Manitoba and Saskatchewan). British Columbia and Ontario come next, followed by the Atlantic provinces, which are not densely populated but have smaller populations.

Table 1 insert here.

Summary Statistics of Inputs and Outputs

Table 2 provides the main statistics on the distribution of inputs and outputs among small hospitals and shows significant variations in all dimensions across these hospitals. The average of total health expenditure is approximately CAD 0.893 million, with a standard deviation of CAD 0.582 million and an interquartile coefficient ((Q3-Q1)/Median) of 0.67 and a range of [30.5 – 1526.8] million. Inpatient expenditure represents approximately half of total expenditure with the average of CAD 0.380 million (SD=0.214) and an interquartile coefficient of 0.60. Small hospitals provided on average 302.2 weighted inpatient cases, with a SD of 220.9 and interquartile coefficient of 0.85 and a range of [16.8 – 1,215.0]. When combining inpatient care and day surgeries, small hospitals provided on average 565.3 weighted cases with a SD of 524.9 and an interquartile coefficient of 1.22 and a range of [16.8, 2675.3]. On average, the number of outpatient care services (ED visits and clinic visits) that small hospitals delivered was 87.4, with an SD of 82.2 and an interquartile coefficient of 1.22 and a range of [1, 457.7]. The average length of stay is less than 1 week (6.6 days), with an SD of 2.5 and an interquartile of 0.47 and a range of [2.5, 14.3].

Table 2 insert here.

Technical Efficiency

Estimated efficiency scores are reported in Table 3. The average output orientation technical efficiency on the acute inpatient care is 58% (SD=15%) meaning that on average, small hospitals in Canada could produce 42% more acute inpatient care using their current endowment. 17 out of 229 small hospitals have an efficiency score above 80% and 67 small hospitals are below 50%. There is significant variation in technical efficiency across provinces: small hospitals in the provinces of Nova Scotia, Ontario and Prince Edward Island are more efficient, at about 65%, than in other provinces.

When the output includes all types of care the overall technical efficiency is 4% less than that based on acute inpatient care only. Nine out of 229 small hospitals have a technical efficiency score above 80% and 89 small hospitals are below 50%. The technical efficiency scores vary from a minimum of 21% to a maximum of 86%. Additionally, small hospitals in the province of Prince Edward Island (66%) are more efficient than that in other provinces. The average technical efficiency in the province of Ontario is 62%. The overall technical efficiency scores in term of all types of service across province ranged between 46%-66%.

Table 3 insert here.

Scale Efficiency

If the output is restricted to acute inpatient care only, the average scale efficiency is as low as 41% and a majority of small hospitals (196 out of 229) have a scale efficiency below 50%. On average, small hospitals in the province of Alberta (32%) and Saskatchewan (44%) are more scale efficient than that in other provinces. The average scale efficiency in the province of Ontario is 41%. The average scale efficiency varies across provinces from 32% - 44 %.

Including all types of services, scale efficiency increases to 66% (SD=19%), meaning that on average, small hospitals should reduce their sizes by 34% without affecting their current production level. In total, majority of hospitals (82%) observed decreasing returns to scale, 35 hospitals increasing returns scale, followed by constant return to scale (6 hospitals). Around one fourth of the small hospitals (55 out of 229) have scale efficiency scores above 80% but about 21% of small hospitals (47 out of 229) are below 50%. The scale efficiency score varies from a minimum of 17% to a maximum of 100%.

The scale efficiency in the province of Nova Scotia (76%) is higher than that in other provinces. On average, the provinces of Alberta, British Columbia and Manitoba have similar scale efficiency (67%). The province of Ontario is the next with the average scale efficiency 61%. The range of the scale efficiency on average across province is between 58% and 76%.

Figure 1 insert here.

Determinants of technical and scale efficiencies

Table 4 reports the summary statistics of the determinants of efficiency in this study and Table 5 presents the results of bootstrapped truncated regressions in which overall technical and scale efficiencies (outputs including all types of service) are regressed against variables that are expected to influence hospital performance.

Table 4 insert here.

With respect to technical efficiency, the percentage of any acute transfer was found to have a positive and statistically significant effect, as expected. It implies that small hospitals would be more efficient if they had the ability to find a larger hospital in a timely manner that could provide suitable treatment for their patients as they need. The coefficient of percentage of discharges to home care was found positive and significant, indicating hospitals with better ability to discharge patients to home care are more efficient. The results also suggest significant associations between context and technical efficiency: hospitals located in higher income health regions have higher technical efficiency. The prevalence of obesity in the health region where a hospital is located is negatively related to its technical efficiency.

One of the key findings on the overall scale efficiency is that small hospitals with a higher percentage of diagnostic tests (cost share) are more likely to be scale efficient. Specifically, one percentage increase of spending on diagnostic tests would lead to an increase in hospital expected scale efficiency score by 1.1%, holding all other variables constant. The percentage of nursing expenses is positively related to scale efficiency, specially, on percentage increase of spending on nursing would lead to an increase in the expected scale efficiency score by approximately 0.3%.

Table 5 insert here.

Conclusion and Discussion

The findings from the first-stage analysis indicate that, on average, the technical and scale efficiencies of all types of service across small hospitals are low (54% and 66%). Small hospitals could increase their outputs (total outpatients visits and hospital discharges) by 46%

with the same budget. In other words, there is scope for treating an extra 177,800 outpatient visits or making 103,800 more RIW weighted inpatient discharge with the existing hospital input endowments. The results of the overall scale efficiency suggest that the hospital size of the small hospitals could be reduced 34% on average for their particular input-output mix, i.e., the Canadian small hospitals would have needed to close approximately 6 beds (the average number of beds is 16) to achieve the optimal scale.

Our second-stage analysis shows that the percentage of acute transfers and of discharge to the home care are significantly correlate with technical efficiency. Thus, policy interventions that create and strengthen transitions and care coordination would reduce inefficiencies and make small and rural hospitals more efficient users of public resources. Even though distance to the nearest urban hospital is certainly a barrier that will not be easy to remove, the development of technology such as telemedicine makes it possible that small hospitals not in close proximity to any urban hospital can still work in network and benefit from services from these large urban hospitals, or collaborate with each other. Telemedicine is not like Health Line which only offers advice, and often requires patients to seek in-person care, but directly connects patients and licensed health care providers online and aims to cut down on in-person visits, making medical care more efficient for both patients and healthcare providers. In rural remote communities, people with limited mobility or limited transportation options could benefit from telemedicine. This study also found that the prevalence of obesity in the region where a hospital is situated has a negative effect on technical efficiency. Policies focusing on improving population health in rural areas would help small hospitals be more efficient.

Results show that small hospitals are scale inefficient at their current input-output mix on average. It is somewhat counter-intuitive because the small hospitals are often perceived as disadvantaged by a lack of economies of scale. However, it is not that surprising because, even though the demand for health care in rural areas is low, driving occupancy rates down, small hospitals must maintain a certain staffing level to be able to provide core service for rural residents and are oversized as a result. It would not be prudent, however, to ameliorate them by down-scaling their health workforce because rural hospitals are often the sole local source for patient care in rural communities. Their potential downsizing or closing might lower health care cost, it may also impact health care access. This suggests that access to timely care when needed could be considered an outcome of small hospitals.

Since it is not realistic to reduce the size of small hospitals, the main area of improvement is in technical efficiency, given a level of size inefficiency. One possible such area of improvement is to decrease administrative expenses that do not contribute to better care. Kaul (2014) advises that multi-hospital network can achieve efficiency, increase scale and lower labor costs by centralizing nonclinical, non-differentiating functions such as human resource, shipping and purchasing. It is encouraged for small hospitals to partner with other small facilities in the same environment and with the same administrative requirement, for instance, in procurement of equipment, supplies and even employee’s benefits. Small hospital could also become more scale efficient by increasing their spending on the diagnostic tests as the percentage of the total hospital spending. Consistent with the point above, instead of each small hospital offering the same limited diagnostic test services, they can invest and offer a broader range of complementary services through a network of small/rural hospital network.

Policy intervention could focus on the measurement of the transfer time, which should also be adjusted by the severity of patients. Most small hospitals have fewer beds and less technology, thus, different from the urban hospitals where acute inpatient care is the primary responsibility, outpatient care, such as emergency care and clinic visits, comprises the first and essential service of small hospitals. However, performance measurements for emergency department care or the clinic visits are not well developed in terms of the quality of care indicators. Schull et al. (2010) developed a consensus on a set of indicators to measure and compare the quality of care in Canadian EDs, including patient satisfaction, ED operation, patient safety, pain management sepsis or infection. However, the majority of these indicators are not available in current provincial or national administrative databases. Especially, none of them reflects the special circumstances of small hospitals or the particular concerns of people living in a rural community. Future research could develop methods to measure the quality of care that reflects not only the main services of the small hospital but also the rural context.

The study reported in this paper has a number of limitations. First, the analysis is based on hospital inputs and outputs data for 2012-2013, not longitudinal data across several years, and we are not able to test for causal effects of determinants of efficiency among small hospitals. We are content with illustrating potential significant associations with efficiency. Second, using the inverse of the length of stay as an indicator of quality of care is not a perfect choice, which only reflects the ability of a small hospital to provide timely treatment to patients. In addition, our study ignores quality for outpatient care. Future research should consider patient satisfaction and the time from arrival to triage for the ED visits. Third, the findings and conclusion are restricted by the availability of data. Future research could target to the small hospitals that perform well on both technical efficiency and scale efficiency. It is necessary to further explore their clinical, management and environment characteristics so as to establish a standard to aim for in trying to improve the technical and scale efficiency of all small hospitals.

Another limitation is that data was not available for the province of Quebec for the study period. According to CIHI, data collection methods and certain access restrictions to Quebec government data preclude CIHI from publishing data from that province on its website. This is unfortunate as Quebec is Canada's second most populated province. Quebec rural hospitals also have better access to CT scanners than other Canadian provinces (74% vs. less than 10%) (Fleet et al., 2013, 2014). Further comparison of efficiency in Quebec rural hospitals with rural hospitals in the rest of Canada will be of interest considering their differential access to in-hospital services.

Funding Source:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References:

Allin, S., Grignon, M., & Wang, L. 2016. "The determinants of efficiency in the Canadian health care system." *Health Economics, Policy and Law*, 11(1), 39-65.

Allin, S., Veillard, J., Wang, L., & Grignon, M. 2015. "How can health system efficiency be improved in Canada?" *Healthcare Policy*, 11(1), 33.

Asmild M, Hollingsworth B, Birch S. 2013. "The scale of hospital production in different settings: one size doesn't fit all." *J Prod Anal*, 40: 197-206.

Baernholdt, M., Jennings, B. M., Merwin, E., & Thornlow, D. 2010. "What does quality care mean to nurses in rural hospitals?" *Journal of Advanced Nursing*, 66(6), 1346-1355.

Banker R.D., Charnes A., Cooper W.W. ,1984. "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis." *Management Science*, 30, pp. 1078-1092.

Bollman, Ray D. and Heather A. Clemenson. 2008. "Structure and Change in Canada's Rural Demography: An Update to 2006." *Rural and Small Town Canada Analysis Bulletin* Vol. 7, No. 7 (Ottawa: Statistics Canada, Catalogue no. 21-006-XIE).

Brasel, K. J., Lim, H. J., Nirula, R., & Weigelt, J. A. 2007. "Length of stay: an appropriate quality measure?" *Archives of Surgery*, *142*(5), 461-466.

Canadian Institute for Health Information (CIHI). 2017. DAD Resource Intensity Weights and Expected length of Stay. Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2016. "National Health Expenditure Trends, 1975 to 2016." Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2016. "Hospital Standardized Mortality Ratio: Technical Notes." Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI). 2016. "Health Indicators 2016." Ottawa, ON: CIHI.

Canadian Institution for Health Information (CIHI), 2013, Hospital Births in Canada: A Focus on Women Living in Rural and Remote Areas. Ottawa: Canadian Institute for Health Information, 2013

Canadian Institute for Health Information (CIHI). 2013. "MIS Standards 2013". Ottawa, ON: CIHI.

Charnes, A., Cooper, W. W., & Rhodes, E. 1981. "Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through." *Management science*, 27(6), 668-697.

Chilingerian J.A., Sherman H.D. 2011 Health-Care Applications: From Hospitals to Physicians, from Productive Efficiency to Quality Frontiers. In: Cooper W., Seiford L., Zhu J. (eds) Handbook on Data Envelopment Analysis. International Series in Operations Research & Management Science, vol 164. Springer, Boston, MA

Chowdhury H, Zelenyuk V 2016. "Performance of hospital services in Ontario: DEA with truncated regression approach." Omega. 2016;63:111–22.

Chowdhury, H., Wodchis, W., and Laporte, A., 2011. "Efficiency and technological change in health care services in Ontario: an application of Malmquist productivity index with bootstrapping." *International Journal of Productivity and Performance Management*, 60(7): 721-745.

Fare, R., Färe, R., Grosskopf, S., & Lovell, C. K. 1994, Production frontiers. Cambridge university press.

Farrell M. and Fredrika Scarth. 2017. "Health system funding reform-new direction." Present at the health Achieve. http://www.healthachieve.com

Fleet R, Archambault P, Legare F, Chauny JM, Levesque JF, Ouimet M, et al. 2013. "Portrait of rural emergency departments in Quebec and utilisation of the Quebec Emergency Department Management Guide: a study protocol." *BMJ Open*.3(4).

Fleet R, Audette LD, Marcoux J, Villa J, Archambault P, Poitras J. 2014. "Comparison of access to services in rural emergency departments in Quebec and British Columbia." *CJEM*. 16(6):437–48. pmid:25358275.

Fragkiadakis, G., Doumpos, M., Zopounidis, C., & Germain, C. 2016. "Operational and economic efficiency analysis of public hospitals in Greece." *Annals of Operations Research*, 247(2), 787-806.

Gonzalez, K., Morganti, S. Bauhoff, J. C. Blanchard et al. 2013. "The Evolving Role of Emergency Departments in the United States." RAND Research Report, 2013.

Guru, V., Anderson, G. M., Fremes, S. E., O'Connor, G. T., Grover, F. L., Tu, J. V., & Consensus, C. C. S. Q. I. 2005. "The identification and development of Canadian coronary artery bypass graft surgery quality indicators." *The Journal of thoracic and cardiovascular surgery*, 130(5), 1257-e1.

Gruca, T. S., and D. Nath. 2001. "The technical efficiency of hospitals under a single payer system: the case of Ontario community hospitals." *Health care management science* 4(2): 91-101.

Jarvis, P. R. E. 2016. "Improving emergency department patient flow." *Clinical and experimental emergency medicine*, 3(2), 63.

Kalhor R, Amini S, Sokhanvar M, Lotfi F, Sharifi M, Kakemam E.2016. “Factors affecting the technical efficiency of general hospitals in Iran: data envelopment analysis.” *J Egypt Public Health Assoc*. 2016;91:20–5.

Kaul A. 2014. “A Belt-Tightening Regimen for Hospital Systems. Hospitals & Health Networks Daily”. http://www.hhnmag.com/articles/4067-a-belttightening-regimen-for-hospital-systems. Accessed January 27, 2017.

Mujasi PN, Asbu EZ, Puig-Junoy J. 2016. “How efficient are referral hospitals in Uganda? A data envelopment analysis and tobit regression approach.” *BMC Health Serv. Res*.16:230.

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

OECD. 2011. “OECD Guidance, Competition Assessment Toolkit.” OECD Vol 2, Version 2.0.

Rechel, B., A. Džakula, G. Duran, N. Fattore, M. Edwards, M. Grignon, T. Haas, G. Habicht, P. Marchildon, A. Moreno, W. Ricciardi, L. Vaughan, and T. Anderson Smith. 2016. “Hospitals in rural or remote areas: a review of policies in 8 high-income countries.” *Health Policy* 120(7): 758–769.

Stewart, D., Black, C., Martens, P., Peterson, S., & Friesen, D. 2000, “Assessing the performance of rural and northern hospitals in Manitoba: A first look”, *Winnipeg: Manitoba Centre for Health Policy and Evaluation*.

Simar, L., and P.W. Wilson. 1998. “Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models.” *Management Science* 44(1):49-61.

Simar, L., and P.W. Wilson. 2007. “Estimation and inference in two-stage, semi-parametric models of production processes.” *Journal of Econometrics*,136(1): 31-64.

Schull, M. J., Banda, H., Kathyola, D., Fairall, L., Martiniuk, A., Burciul, B., ... & Mondiwa, M. 2010. “Strengthening health human resources and improving clinical outcomes through an integrated guideline and educational outreach in resource-poor settings: a cluster-randomized trial.” *Trials*, *11*(1), 118.

Southard PA, Hedges JR, Hunter JG, Ungerleider RM 2005. “Impact of a transfer center on interhospital referrals and transfers to a tertiary care center.” *Acad Emerg Med*. 2005 Jul;12(7):653–657.

Sutherland, J., Crump, R., Repin, N. and Hellsten, E., 2013. Paying for hospital services: a hard look at the options.

Sultan, W. I., & Crispim, J. 2018. "Measuring the efficiency of Palestinian public hospitals during 2010–2015: an application of a two-stage DEA method." *BMC health services research*, 18(1), 381.

Wang L., Grignon M. Perry S., Chen X., Ytsma A., Allin S., and Gapanenko K. 2018. "The Determinants of the Technical Efficiency of Acute Inpatient Care in Canada." *Health Service Research*, 53(6):4829-4847. doi: 10.1111/1475-6773.12861.

Wilson, P. W. 2008. "FEAR 1.0: A Software Package for Frontier Efficiency Analysis with R." *Socio-Economic Planning Sciences* 42, 247--254.

Xenos, P., Yfantopoulos, J., Nektarios, M., Polyzos, N., Tinios, P., & Constantopoulos, A. 2017. "Efficiency and productivity assessment of public hospitals in Greece during the crisis period 2009–2012." *Cost Effectiveness and Resource Allocation*, 15(1), 6.

Table 1: The distribution of the small hospitals in Canada

Province	**# of hospitals**	**%**
New Brunswick	0	0
Prince Edward Island	5	2
Newfoundland	14	6
Nova Scotia	19	8
Ontario	31	14
British Columbia	32	14
Saskatchewan	39	17
Manitoba	41	18
Alberta	48	21
Total	229	

Table 2: Summary Statistics of the inputs and outputs

	Mean	SD	Range	
			Min.	Max.
Input(1,000')				
Inpatient cost	453.7	267.8	30.5	1526.8
Total cost	1093.3	736.0	195.2	4981.4
Output				
Inpatient RIW (100')	393.5	313.7	16.8	1956.6
Inpatient +Day Surgery RIW (100')	710.9	648.2	16.8	3170.0
Emergency visits + Clinic Visits (100')	121.8	113.5	0.01	639.8
100/length of stay	17.4	6.4	7.0	46.5

Note: Inpatient cost – not including the capital, cost in the mental and rehabilitation inpatient care; Total cost – not including the capital, cost in the mental and rehabilitation inpatient care.

Table 3: The technical and scale efficiency of small hospitals in Canada

	# of hospitals (%)		Summary Statistics			
Efficiency	> 80%	< 50%	Mean	Std. dev.	Min.	Max.
Inpatient services						
Technical	17(7.4%)	67(29.3%)	0.58	0.15	0.18	0.93
Scale	6(2.6%)	196(85.6%)	0.41	0.13	0.11	1.00
Overall services						
Technical	9(3.9%)	89(38.9%)	0.54	0.14	0.21	0.86
Scale	55(24.0%)	47(20.5%)	0.66	0.19	0.17	1.00

Table 4: Summary statistics of the determines of the efficiency

Variable	%	Std. Dev.	Range (%)	
			Min	Max
% Long stay patients	11.98	6.86	1.28	38.30
% Patient transferred (in or out)	21.70	8.04	0.94	66.67
% Patients discharged to home with care	10.11	9.19	0.00	43.85
% administration expenses	30.63	6.59	8.79	67.79
% diagnostic tests	13.74	4.97	0.50	30.80
% nursing expense (inpatient)	35.08	12.46	8.69	68.19
Average income at the neighborhood ($1,000) *M*	64.86	32.02	20.96	125.03
% neighborhood obsess	23.04	4.49	7.42	35.80

Table 5: The double bootstrap Tobit regression on the technical and scale efficiency based on overall service output

	Technical Efficiency			**Scale Efficiency**		
	Coefficient	Std. Err.	95% CI	Coefficient	Std. Err.	95% CI
% Long stay patients	-0.199	0.125	[-0.446 , 0.048]	-0.083	0.211	[-0.499 , 0.333]
% Patient transferred (in or out)	**0.977*** **	0.107	**[0.766 , 1.188]**	0.097	0.18	[-0.259 , 0.452]
% Patients discharged to home with care	**0.260*** **	0.081	**[0.100 , 0.420]**	0.067	0.137	[-0.202 , 0.336]
% administration expenses	**-0.258** **	0.121	**[-0.496 , -0.021]**	-0.201	0.203	[-0.602 , 0.199]
% diagnostic tests	0.18	0.168	[-0.15 , 0.51]	**1.051*** **	0.282	**[0.495 , 1.606]**
% nursing expense (inpatient)	**-0.192*** **	0.062	**[-0.313 , -0.071]**	**0.276*** **	0.104	**[0.072 , 0.48]**
Average income at the neighborhood ($1,000) *M*	**0.043*** **	0.014	**[0.015 , 0.071]**	**-0.042* **	0.024	**[-0.089 , 0.006]**
% neighborhood obsess	**-0.005*** **	0.002	**[-0.008 , -0.002]**	0.002	0.003	[-0.004 , 0.008]
R square	47.9%			11.40%		

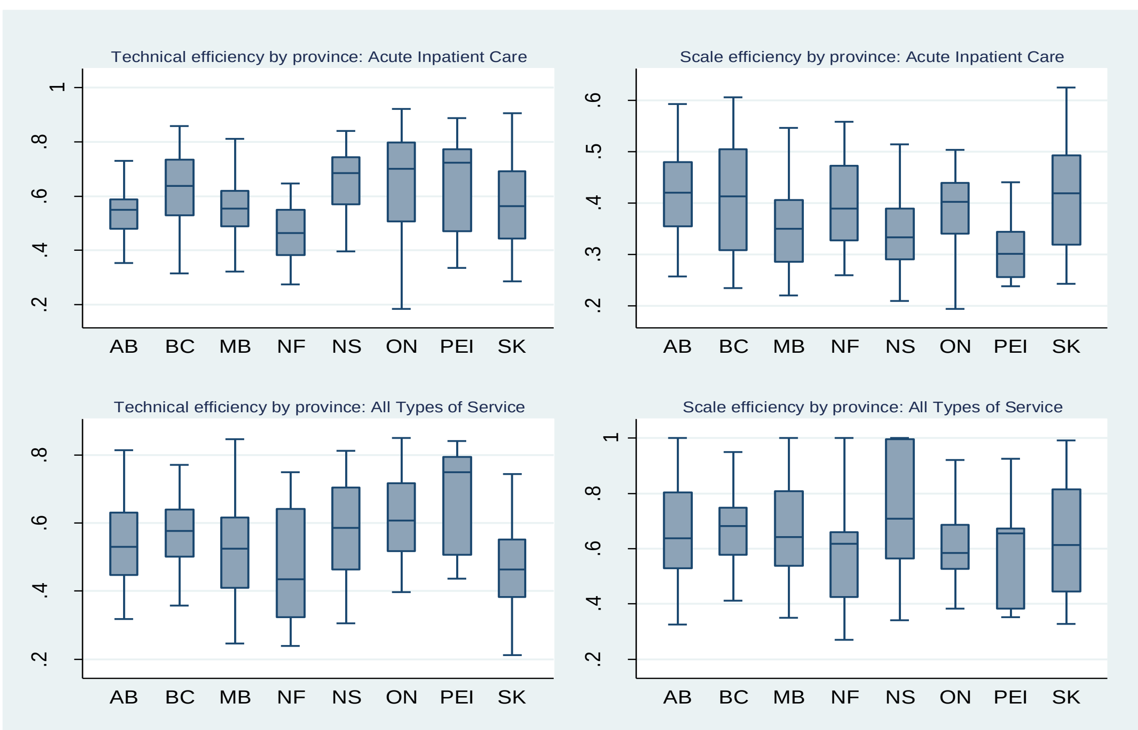

Figure 1: Technical and Scale efficiency by province

CHAPTER V: SUMMARY OF NOVEL CONTRIBUTIONS AND FUTURE DIRECTIONS

Major findings and future directions

This thesis addresses some novel questions surrounding the efficiency of Canadian hospitals. This work provides several original contributions and I will highlight them by chapter and discuss future research directions.

Chapter 2: The determinant of efficiency of acute inpatient care

The main findings of this paper demonstrated that the average of technical efficiency of acute inpatient care in Canada (except for the province of the QC and territories) is 76% by applying a non-parametric DEA approach to teaching, large and medium hospitals. In other words, the hospital output would be increased by 24 percent with the same resources by eliminating inefficiency. The determinants of acute inpatient care efficiency are summarized into three dimensions by using a principle component analysis: Advantaged environment, higher proportion of non-complex patients, health system cooperation and availability of post-acute care and the poor quality of overall quality of care.

Specifically, highly efficient teaching hospitals are associated with advantaged environments, such as the percent of patients living in urban areas and the higher education of the population in the region where the hospital is located. Among large hospitals, increasing the proportion of non-complex patients, such as the percent of the newborn cases and the percent of long stay patients, is associated with improved hospital technical efficiency. Greater cooperation within the health system and the availability of post- acute care beds is associated with improved efficiency for both large and medium hospitals. The technical efficiency of medium hospitals is positively associated with the advantaged environment factors, such as the population's higher education and the patients living in the urban, as well as the higher proportion of non-complex patients.

Generally, the main drivers of efficiency of acute inpatient care vary by hospital peer groups. Thus the results provide different policy and managerial implications for teaching, large and medium hospitals to achieve efficiency gains.

Chapter 3 The spatial interdependence of acute inpatient care in Canada

In this novel contribution, a DEA with bootstrap methodology is employed to estimate the bias-corrected technical efficiency on a sample of 179 acute hospitals. The novelty is to employ a spatial lag model in a second stage to estimate the spillover effect of technical efficiency while controlling for potential managerial and environmental factors. I find a substantial and significantly positive spatial spillover effect on the efficiency of acute inpatient care within catchment areas defined using 60km to 240KM as the upper limit of influence and the magnitude generally declines with distance. The spatial analysis also identifies hospital size (the number of beds), the proportion of acute transfers, and the percent of patient transfers to home care as the main drivers of efficiency of acute inpatient care in Canada. Further work is however suggested to use the direct indicator of the hospital interaction as the spatial matrix, such as the quantity of acute transfers between hospitals, to estimate the peer effects on the hospital efficiency. Hospitals in Canada chose cooperative strategies rather than competition to improve their technical efficiency, through transfers of patients on the basis of need. This may be linked to the main mode of payment (global budget), even though many provinces are now moving toward activity-based funding.

Chapter 4 The technical and scale efficiency of small hospitals in Canada

The objectives of the study are to estimate the relative technical and scale efficiencies of 229 small and remote hospitals in Canada using DEA, as well as to estimate the impact of institutional contextual/ environmental variables on hospital technical and scale efficiency by using the double bootstrapping Tobit regression.

The study is found that small hospitals achieve a lower level of either technical or scale efficiency in Canada than large and urban hospitals. The average technical efficiency score for all types of service is 54%, indicating small hospitals could produce 46% more health service outputs using their current input endowment. The scale efficiency based on all types of service was 66%, meaning that on average, the small hospitals could reduce their size by 34% without affecting their current output levels. In the second stage, double bootstrap Tobit regression were estimated in which measures of overall technical and scale efficiencies were regressed against set of explanatory variables respectively. The results show that improving the ability to discharge patients to home care and enforcing the corporation with other acute large, teaching or medium hospitals to transfer complicated patients out for the treatments and transfer in the patients for recovery would increase the technical efficiency of small hospitals. In addition, we find the percentage of the administration expenses and nursing expenses as a proportion of total costs are negatively associated with the technical efficiency. The contextual factors, such as the average income and percentage of the obesity rate in the hospital neighborhood are all significant determinants of technical efficiency. In addition, we also found that small hospitals with more spending on the diagnostic tests and nursing tend to be more scale efficient.

The policy maker would aim to enforce the collaboration with the urban hospitals and the post-acute home care, as well as to develop context-appropriate strategies for supporting hospitals with low technical efficiency to improve their service and thereby better address unmet needs for hospital services in Canada

Overall Conclusion

Inpatient acute care is the most resource intensive health care service. To reduce hospital costs, many jurisdictions have been coordinating services and increasing alternatives to inpatient care such as day-surgery, specialized clinics, diagnostic imaging centers, home care and long-term care. Hospitals interact with each other and other health care organizations and providers. They share expensive equipment; provide specialized care for the region or province; redistribute services to address a high demand for certain procedures in the region. Hospitals encourage community care by integrating systems to exchange patient information facilitating faster and better patient care. There is a push for increased collaboration and coordination between hospitals and the community particularly primary care doctors, long-term care organizations and home care. Efficiencies can be found in consolidation, integration, and team care. Hospital's efficiency might be improved by availability of these services in the region and province as they can discharge their patients faster and save the resource for the need of other acute care.