COMMENTARY

The Nordic health care systems: Most similar comparative research?

KJELD MØLLER PEDERSEN¹,

¹Department of Business and Economics, University of Southern Denmark, Denmark

Abstract: Five comparative articles on the Nordic health care systems are discussed stressing the need for comparative methodology and better understanding of the volume-outcome literature underlying much of the centralization of hospital systems.

JEL classification: I11, I14, I18

Key words: volume-outcome, quality, centralization, comparative method

1 Introduction

Comparison is the essence of much science. Comparison functions as a mirror – and occasionally as a magic mirror. Researchers know, however, how difficult it is meaningfully to compare national health systems because of different institutional setups: From financing over regulatory schemes to the role of central and decentral political bodies. Therefore, it is always fraught with danger to transfer and compare findings from one health system to another, and if done, it requires good knowledge of system related institutional details. This is an important common observation across the five papers commented on here where two to five Nordic countries are compared.

I start with a discussion of the much-neglected issue of comparative methods and institutional details followed by comments and suggestions on the individual papers with particular focus on the volume-outcome issue that run across several of the articles along with multi level analysis.

2 Comparison and institutional detail

The Nordic countries in many respects have rather similar health systems, and hence invite comparison despite the fact that the greatest similarity may be the institutional context in which the health care sector is embedded (Lyttkens et al., 2016) and not the detailed institutional organization.

In the terminology of comparative research, the Nordic health systems are most similar system, but still challenging to compare. The Nordic systems can be likened to a living laboratory that invites comparison, in part due to reasonable good comparative data. Hence, results from comparative Nordic studies should always be welcomed.

In the language of comparative research (Przeworski and Teune, 1970; Anckar, 2008) which ought to be the starting point for comparative methodology, Nordic health care systems are most similar, i.e. we compare relatively similar cases that only differ in the
dependent variable, typically an outcome variable like performance or mortality, on the assumption that this would make it easier to find those independent variables which explain the presence/absence of the dependent variable when many of the potentially relevant independent variables are rather similar at the starting point. Most similar can be contrasted to most different systems.

It is amazing to find how cavalierly institutional setup is approached even by experienced and internationally known health economists, e.g. when US findings on for instance competition uncritically is transferred to for instance a north European setting. An example being Enthoven’s ideas about ‘managed competition’ (Enthoven, 1985; Enthoven, 1988; Enthoven, 2000) that too uncritically were transferred to Margaret Thatcher’s and John Major’s UK health care reform around 1990. How should American experience and setting meaningfully be transferred to the radically different NHS in the UK? Or when Kaiser Permanente was compared to the NHS (Richard et al., 2002) or Denmark (Schiøtz et al., 2011).

Joseph Newhouse in his 2002-book on Pricing the Priceless – a Health Care Conundrum (Newhouse, 2002) apologized to his non-American readers that his examples overwhelmingly drew on US. “The non-American reader may sometimes have to work a bit to relate the examples to the institutions of his or her country, but I believe this will almost always be possible”, p. 6 (Newhouse, 2002). Anyone who has read the book will see that this is a gross understatement. Furthermore, he noted that he would “mostly abstract from … institutional differences among countries. That does not mean the differences are unimportant … but … theories and models are typically sufficiently general that they will apply to many institutional settings”. In some cases, this may be true, e.g. Grossman’s model for the demand for health. In other cases, it does not hold. Consider the interest in value based health care in for instance Sweden and Denmark (Pedersen, 2017) fueled by Porter’s book on Redefining health care - Creating Value-based Competition on Results (Porter, 2014). Apart from Michael Porter not having a background in health economics, the setting for his and Teisberg’s book is the US, very US. For instance, the focus on bundled payment (Porter and Kaplan, 2016; Agency for Healthcare Research and Quality, 2012) has a different meaning in the US than in for instance Denmark, the core issue being what DRG rates include. In the US they for instance do not include physician costs while this is the case in Denmark, and ‘bundling’ in the US often just imply that physician costs are now are ‘bundled’ into the DRG-rate. In fairness to Porter’s ideal world bundled payment ideally is tied to patient outcome. However, the point is that transferring models and modes of thinking from one country to another is fraught with dangers.

Models and methods for comparative health services/systems are not very well developed (Marmor and Wendt, 2012; Papanicolas and Smith, 2013). Hopefully the studies in this edition of the Nordic Journal of Health Economics is a first step in the direction of rectifying this situation. The articles as such do not address methodological issues directly but presupposes relevant comparison, e.g. Helsinki and Oslo.

Many ‘comparative’ studies are worthless at least from a policy point of view despite statements to the contrary. This is not true, however, of the articles in this edition of the journal even though one should be careful not to take all results at face value. However, take the WHO 2000 World Health Report (WHO, 2000; Pedersen, 2002) and the econometric work on Comparative efficiency of national health systems (Evans et al., 2001). The intention was to improve the evidence base for health policy by devising a method to measure and monitor the performance of health systems by estimating the relation between levels of population health and the inputs used to produce health. A laudable ambition – but to do it for 191 countries and develop a ranking does not make sense (disregarding econometric issues (Pedersen, 2002)). Ranking presupposes an assumption
about a reasonable degree of comparability, e.g. Andorra, Jamaica, Japan and Saudi Arabia ranked 7th to 10th respectively (Evans et al., 2001) along with possible explanations which also require institutional details that are not available (Schütte et al., 2018). In 2010 OECD published a working paper providing institutional details for 29 countries (Paris et al., 2010). One of the objectives was to develop a limited set of quantitative indicators designed to capture the main characteristics of health systems. These indicators were intended to assess the role of health institutions and policies on health systems efficiency. This is the tedious and hard work needed to make sense of econometric comparative analyses, and unfortunately all too often neglected leading to misleading or too hasty policy recommendations.

3 The articles

The five papers considered here, table 1, are all empirical in nature and illustrate among other things how careful one must be to work with cross-country data. Apart from the paper on inequality in health care (Christiansen et al., 2018) they are based on register data where for instance a separate and uniform DRG-coding has been done to achieve comparability. There is no doubt that the EuroHOPE - European Health Care Outcomes, Performance and Efficiency (www.eurohope.info) - that evaluates the performance of European health care systems in terms of outcomes, quality, use of resources and costs has been a driving force. Several of the authors have been involved in EuroHOPE-project on Nordic Hospitals, AMI or Hip fractures and draw on data and valuable experience from these projects. A 2015 supplement to Health Economics contained a number of studies from the EuroHOPE project. All of the articles include Nordic co-authors.

The articles are not descriptive but attempt to answer more less clearly stated hypothesizes, e.g. “test if quality-related economies of scale existed in Nordic hospital production” (Kittelsen et al., 2018), but are careful not to claim causal relationships, however, using guarded language: “…we are aware that this cross sectional relationship may not necessarily imply a causal relationship” (Christiansen et al., 2018).

Some of them also has an ambition to be policy relevant. For instance, centralized structure vs less decentralized and the effects of coordination with a clear understanding of the need for relevant institutional details. “Policy-relevant conclusions and implications require more knowledge about the background behind the variations in treatment costs between countries” (Iversen and Häkkinen, 2018) and in the article on acute coronary syndrome, ACS, it is noted that “A further research topic of relevance to policy makers would be to investigate the cost-effectiveness of the alternative organizational forms. If costs do not differ, the organizational form does not matter, but if the total costs of either system including referrals exceeds the other, policy makers should carefully evaluate the options.”(Moger et al., 2018)

4 Comments on the individual papers

4.1 Inequality

The article on inequality demonstrates a general challenge in comparative studies, namely to explain results (outcomes): Mean health and share of the population with good or very good health is higher in the Nordic countries and the UK compared to Germany which may, the authors note, support a tentative hypothesis that centrally initiated public health activities matter. However, life expectancy in Germany is similar to what is found in the other countries with centrally initiated public health activities.
One thing is to measure outcome, another is to explain differences or lack thereof. Christiansen et al. (2018) introduce measures of public health noting that each of the Nordic countries have introduced centrally initiated comprehensive public health programs to increase health and reduce socio-economic inequalities in health. Is it possible to link public health initiatives to measures of inequality used hence providing possible explanations? They find it difficult to do stringently and essentially just hints at a possible relationship. One of the problems is that for inequality measurement they use data on individual level, whereas public health initiatives and lifestyle data are on country level. One approach to

<table>
<thead>
<tr>
<th>Table 1: Overview of the articles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Healthcare, health and inequality in health in the Nordic countries by Christiansen et al. (2018)</td>
</tr>
<tr>
<td><strong>Scale and quality</strong> in Nordic hospitals by Kittelsen et al. (2018)</td>
</tr>
<tr>
<td>Performance comparison of <strong>hip fracture</strong> pathways in two capital cities: Associations with level and change of integration by Häkkinen et al. (2018)</td>
</tr>
<tr>
<td>Comparative treatment costs for patients with <strong>acute myocardial infarction</strong> between Finland and Norway by Iversen and Häkkinen (2018)</td>
</tr>
<tr>
<td>Higher mortality among <strong>ACS</strong> patients in Finland than in Norway: Do differences in acute services and scale effects in hospital treatment explain the variation? by Moger et al. (2018)</td>
</tr>
</tbody>
</table>
reconcile this may be to look at it from the perspective of multilevel models – where level one is individual observations and level two is the country level (public health initiatives/lifestyle variables).

Multilevel models are increasingly used within the social sciences because they allow the possibility of modelling outcomes like inequality as a function of explanatory variables at different levels as just mentioned. Using data from the European Social Survey, Eikemo et al. (2008) explored self-perceived health. The study showed that almost 90% of the variation in health was attributable to the individual-level, while approximately 10% was associated with national welfare state characteristics.

4.2 Scale and quality
The main aim of the article on Scale and Quality in Nordic hospitals is to test if quality-related economies of scale existed in Nordic hospital production. The Nordic EuroHOPE data were re-used for testing whether controlling for quality increases estimates of the elasticity of scale. There could be important policy implications if increased volume allows for reduced unit costs with the same quality, or alternatively allows for increased quality at unchanged unit costs. It is implicitly assumed that the optimum volume for economies of scale and volume-outcome coincide. This is more than doubtful and based on a somewhat incomplete reading of the rather voluminous outcome-volume literature (Spaulding et al., 2006; Urbach and Baxter, 2004) that is included by only one reference in the article.

The old adage that ‘practice makes perfect’ lends an intuitive reasoning to the relationship between higher operator and hospital procedural volumes and consequent better clinical outcomes (quality). It was first explored by one of the health economic pioneers, Harold Luft, in a pioneering article from 1979 (Luft et al., 1979) asking whether operations should be regionalized and followed up a few years later 1987 (Luft et al., 1987; Hughes et al., 1987). One possible version of the relationship is seen in the lower part of diagram 1 below. The thinking behind the volume-outcome (volume-quality) lies behind much (centralization) planning in the hospital sector. I have explored this at length in chapter 12 of ‘Health Policy’ (Pedersen, 2005), including the relationship between economics of scale and outcome volume, figure 1, where optimum volumes from the economies of scale and volume-outcome perspective have here been designed not to coincide. They are most likely not to do so because they are driven by two different logics: Economics and professional practice (‘practice makes perfect’).

The paper on scale and quality would have gained considerably by trying to include important points from the outcome-volume literature and not by merely adding quality variables to the regression analysis on scale economies (Urbach and Austin, 2005; Urbach and Baxter, 2004).

4.3 Mortality among ACS patients
This study looks at whether a decentralized structure with reduced emergency preparedness and small scale production in Finland vs. a centralized structure with large PCI departments performing acute services 24/7 in Norway, explain the country differences in mortality. The authors were not able to explain the mortality differences from the organizational variables. Rather, the explanation seems to have other sources.

The issue here might to advantage have been framed within the volume-outcome framework providing at one and the same time a possibly explanation of centralization and the different outcomes in more or less centralized hospital structures (Badheka et al., 2016; Wouters et al., 2009). Furthermore, multi-level analysis probably could have been used to
advantage (Urbach and Austin, 2005; Urbach and Baxter, 2004), e.g. experience of individual surgeons vs. hospital level issue, like centralization. However, data may not have been available for this undertaking.

4.4 Hip fracture pathways
Better integrated and coordinated health care is a common challenge in the Nordic health systems. One issue is how to achieve this organizationally, e.g. the Norwegian Coordination Reform in 2012. Another issue is how to measure possible effects of better integration. The paper on hip fracture pathways aim to analyze the associations between different modes of integration and performance indicators. Performance indicators in Oslo and Helsinki were compared at baseline before the Norwegian Coordination Reform, i.e., 2009–2011, and trends were described and analyzed by difference-in-difference methods.

The article raises the question of (in)sufficient institutional detail – a detailed knowledge that the reader does not have readily available. On the positive side is the fact that there is a separate section on institutional settings, but on the negative side there is not sufficient detail on why for instance the different economic incentives in the two cities are sufficiently different to explain performance details like length of stay, e.g. how different are the bed blocking fees in Helsinki and Oslo respectively. There are too few details to justify the hypothesis that there is an inverse relationship between LOS in hospitals and LOS in institutional care following an acute hospital stay, i.e., a short hospital stay in the acute phase will be followed by a comparatively longer rehabilitation period outside the hospital and vice versa. It also seems that the authors believe that their econometric work is a test of the effect of the Coordination reform in Oslo, e.g. a heading like ‘evaluation of the coordination reform’ where we have only been told about two economic incentives: Bed blocking fee and municipal co-financing, but where the latter is not relevant for hip-patients, hence leaving bed-blocking fees as the driving force. I know that there are more in the coordination reform than that, but this is not explicated. Often the econometric models do not capture sufficient institutional details – in part because data is not readily available for econometric work.
4.5 Comparative treatment costs for AMI

The point of departure for the AMI-article is that Finland has higher hospital costs per patient than Norway for the first hospital episode, while Norway has higher costs during the first year after the initial admission. In particular, the focus is on how variables related to the organization of acute care may influence treatment costs.

Variations in the organization of health care for AMI patients between Finland and Norway are mainly related to more centralized emergency care for patients with AMI in Norway compared with Finland. Centralization implies longer travel distances, which may in turn imply longer hospital stays to prevent post-discharge complications. On the other hand, centralization may imply smaller costs per patient due to the potentially positive returns to scale of the specialized treatment for AMI patients. Further, due to centralization, Norwegian patients are treated in higher-volume hospitals than Finnish patients.

One result reported is that there was no volume effect, while the distance from home to hospital increases hospital costs at a declining scale. This takes us back to figure 1 and the need to distinguish carefully between economies of scale and volume-outcome. It is noted that the treatment outcome in terms of mortality seems to matter a lot for costs differences and that more attention should be paid to revealing the mechanisms behind the relationship between hospital costs and mortality. One possible approach is to focus on the volume-outcome issues.

5 Concluding observations

The papers commented on here are empirical based on good (register) data and of course use well-known econometric technique competently. They are good papers based on the usual academic standards. Cross-country comparisons are important, in particular when we are looking at rather similar health systems like the Nordic countries. Great strides have been made regarding data comparability - in part due the to EuroHOPE-project. This is used to advantage in the articles discussed here.

However, it seems relevant to raise issues that ought to be pursued in future studies: 1. More focus on comparative methodology and institutional detail, and 2. Maybe too much ‘measurement without theory’.

The first issue has been developed above whereas the issue of ‘measurement without theory’ need more development. The heading is taken from an old article (Koopmans, 1947) by Nobel prizewinner Tjalling Koopmans, and has more recently been revisited by Wolpin (Rust, 2014; Wolpin, 2013) in his book on The Limits of Inference Without Theory. Koopmans article grew out of a review of a book on Measuring Business Cycles and the very ‘empiricist approach’ taken, i.e. no theory guided the work. Kenneth Wolpin – revisiting the issue more than 60 years later - found that the absence of theory in inferential work that addresses microeconomic issues was pervasive. That theory is unnecessary for inference is exemplified by the expression “let the data speak for themselves”. This approach is often called “reduced form.” I am not saying that this is the case for the articles reviewed here, but I am worried at a more general level. I see tendencies in the empiricist direction in much empirical work due to a) available and reasonable comparable data, b) easily accessible econometric programs, and c) good econometricians. There is a tendency to neglect a more theory based development of hypotheses. I am not asking for hypotheses developed from comparative-static models. However, much too often hypotheses are developed offhanded by hinting at theory or theories and then proceed with the econometrics. This in combination with lacking institutional details should raise concern.
References


© 2019 by the author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).