Eur J Health Econ (2016) 17:1011–1026
DOI 10.1007/s10198-015-0743-4

ORIGINAL PAPER

Is higher nursing home quality more costly?

L. Di Giorgio[1,2] · **M. Filippini**[2,3] · **G. Masiero**[2,4]

Received: 25 July 2014 / Accepted: 21 October 2015 / Published online: 26 November 2015

Abstract Widespread issues regarding quality in nursing homes call for an improved understanding of the relationship with costs. This relationship may differ in European countries, where care is mainly delivered by nonprofit providers. In accordance with the economic theory of production, we estimate a total cost function for nursing home services using data from 45 nursing homes in Switzerland between 2006 and 2010. Quality is measured by means of clinical indicators regarding process and outcome derived from the minimum data set. We consider both composite and single quality indicators. Contrary to most previous studies, we use panel data and control for omitted variables bias. This allows us to capture features specific to nursing homes that may explain differences in structural quality or cost levels. Additional analysis is provided to address simultaneity bias using an instrumental variable approach. We find evidence that poor levels of quality regarding outcome, as measured by the prevalence of severe pain and weight loss, lead to higher costs. This may have important implications for the design of payment schemes for nursing homes.

Keywords Nursing home · Costs · Nonprofit · Single quality indicators · Composite quality indicators · Cost-quality tradeoff · Process quality · Outcome quality · Structure quality

JEL Classification I10 · L3

Introduction

In many European countries, the delivery of nursing home services occurs mainly through public and private nonprofit providers, which may be financed by different payment schemes. Widespread concerns about nursing home quality have played an important role in the recent political discussions about increasing funding [1]. The underlying assumption is that more resources are needed to boost quality levels. However, little evidence is available to support this hypothesis in the nonprofit sector and for different payment systems. The aim of this analysis is to provide evidence on the relationship between quality and costs in one European country, Switzerland. This evidence may help to inform policymakers in countries that share the same characteristics of the Swiss nursing home sector.

Ideally, we would want to investigate this relationship by pooling data from different European countries under the same delivery system. However, for our analysis, we choose to focus on one country since access to data represents a barrier, and country and regional differences in regulation and definition of nursing home services are likely to confound our findings.

We thank Andrew Street for helpful comments and invaluable advice during our stay with the Policy team at the Center for Health Economics at the University of York. Also, we thank the Swiss National Science Foundation for financial support to the project. Any errors are the authors' responsibility.

✉ G. Masiero
giuliano.masiero@unibg.it

[1] Institute for Health Metrics and Evaluation (IHME), University of Washington, Washington, United States

[2] Institute of Economics (IdEP), Università della Svizzera italiana (USI), Lugano, Switzerland

[3] Department of Management, Technology and Economics, ETH, Zurich, Switzerland

[4] Department of Management, Information and Production Engineering (DIGIP), University of Bergamo, Bergamo, Italy

Switzerland is a federal state with 26 independent cantons which are granted extensive autonomy in the provision of long-term care services and other social services. Since the high level of autonomy led to 26 heterogenous systems, we limit the analysis to one homogeneous region (Canton Ticino). In Canton Ticino, nursing home services are mainly provided by public (46.5 %) and private nonprofit (48.5 %) nursing homes (NHs), while for-profit providers represent a small minority (5 %). However, private NHs are excluded from the cantonal administration and are not required to share their data. The provision of nursing home care is further decentralized at the local level (municipalities), and elderly people are commonly assigned to the NH in the community of residence, virtually excluding competition and patient self-selection. Prices and some aspects of quality are regulated at the cantonal level. In particular, quality is regulated in terms of structural elements and staffing levels.[1] Capital costs are covered through a retrospective payment system, while operating costs are subject to global budget payment. The global budget payment system replaced the previously-in-force cost reimbursement system in 2006, in order to increase transparency and efficiency in the sector. Consumer fees finance part of the system and are a function of residents' wealth and income (pension payment).

A positive relationship between costs and quality is generally expected when higher levels of quality can be provided through structural and procedural improvements, such as obtaining more costly equipment or additional staff employment. However, adverse patient outcomes may be costly to treat because they involve additional resource utilization for extra care. The relationship between costs and quality may therefore depend on the dimension considered. Better procedures are expected to increase costs, while prevention of adverse outcomes may actually reduce costs. Recent studies on nursing home costs using clinical quality indicators generally include single indicators of quality in cost analyses, possibly neglecting the relationship between different quality dimensions. Since the correlation between quality indicators is usually low, more effort is needed to understand whether the multidimensional nature of quality is better captured by single or combined quality indicators.

Through this paper, we investigate the relationship between quality and costs in NHs using a cost function approach consistent with the economic theory of health care production. We contribute to the existing literature in two main respects. First, we use panel data models to address omitted variables bias, which is particularly relevant when a subselection of indicators is included in the analysis. Second, we disentangle the impact of process and outcome quality dimensions on costs. This is done using both composite and single measures of quality. To our knowledge, this is the first study providing evidence on the relationship between costs and quality in nonprofit nursing home care using panel data, if we exclude the analysis by Wodchis et al. [3], which does not specify a cost function.

The remainder of the paper is organized as follows: "Quality" defines quality measurement in nursing home care. "Empirical evidence on the impact of quality on costs" reviews previous studies on the relationship between costs and quality. "Model specification and data" describes the data set and discusses the choice of quality indicators and the empirical strategy. Estimation technique and results are presented in "Econometric estimation and results". "Conclusions" provides concluding remarks.

Quality

No universal definition of quality exists in health research. The US Institute of Medicine [4] states that "quality of care is the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge".[2] This definition has significantly influenced the literature on quality and is very much related to the paradigm of quality proposed by Donabedian [9]. His seminal article on the assessment of quality of care represents the foundation of modern quality assessment, providing a framework of reference with guidance validity. Donabedian proposed the so-called structure, process, and outcome (SPO) framework. Structure is defined by the attributes of the setting in which care is provided, such as material resources (e.g., equipment), human resources (e.g., staffing levels), and organizational structure (e.g., payment system). Process refers to the activities of practitioners in giving care, such as making a correct diagnosis and implementing the treatment accordingly. Outcome defines the change in health status of the patient.

Inability to include information about these three dimensions of care is due to measurement deficiencies and limitations in data availability. Recently, the introduction of the resident assessment instrument (RAI) in the United States and some European countries started a comprehensive and multidimensional assessment of all nursing home residents' health status. These data, also called minimum data set (MDS), are used to develop a battery of clinical

[1] For a more detailed description of the Swiss nursing home sector, see [2].

[2] Other well-recognized definitions are provided by the UK Department of Health [5], the Council of Europe [6], and the World Health Organization (WHO) [7]. For a detailed exposition of the most influential and known definitions of quality, see [8].

indicators of quality that meet the taxonomy of the SPO model [10, 11]. As such, they offer a unique tool to measure and compare quality of nursing homes in different domains of care [12].

The success of the SPO paradigm lies in its broad scope, which encompasses older and newer definitions of quality. Table 1 shows how different measures of quality used in the literature fall within the dimensions of the SPO framework. The first three columns include non-clinical and clinical indicators of structure, process, and outcome. The last column comprises consumer- and family-reported indicators. In a recent study, Li et al. [13] show that satisfaction ratings are associated with other common indicators of quality, such as higher nursing staffing levels and fewer citations. The authors also find higher scores in public and private nonprofit nursing homes compared to for-profit nursing homes. Unfortunately, satisfaction ratings are available for only one year in the present study and suffer from lack of variation across facilities (<10 %). For this reason, these data cannot be used.

With the development of quality indicators derived from the RAI, clinical measures of quality regarding process and outcome are now available. However, countries still use different systems to measure quality in the nursing home sector [14], and only a few of them have adopted the RAI.

Previous studies attempt to capture nursing home quality differences mainly using indicators of structure or indirect signals, such as the number of deficiency citations [15], staffing levels [16–18], staff characteristics [19–22], and willingness to take leadership [23]. A recent systematic review by Bostick et al. [24] shows not only evidence of association between higher levels of licensing attained by staff members and quality, but also a significant relationship between staff turnover and quality indicators such as pressure ulcers, weight loss, and functional ability. Some relatively old indicators (non-clinical) are still considered valid and are often combined in empirical studies with clinical quality indicators derived from the RAI.

The advantages and disadvantages of quality indicators based on the SPO model are discussed in [15]. Structural indicators are easy to measure, and data are often available. The disadvantage is that the presence of structural attributes does not imply their best use. Indicators of process are usually easy to interpret as they inform on the provision of a particular treatment. Even in this case, it cannot be determined whether or not the provided treatment is appropriate. Finally, outcome indicators are of natural interest, as they measure the change in patients' health status. The main problem with these indicators is that it is extremely difficult to isolate the effect of care and changes in health, as the latter may be influenced by many uncontrolled factors.

The recent development of clinical quality indicators has improved the measurement of quality, but with some limitations. Firstly, due to the absence of a universally accepted definition of quality, the selection of quality indicators to include in empirical analyses is, to some extent, arbitrary. This is an issue because of the usually low correlation among quality indicators. Indeed, facilities with excellent outcomes in some dimensions may perform poorly in others. The choice of indicators may therefore affect the perception of nursing home quality. Secondly, detection bias occurs if higher-quality nursing homes are more vigilant in looking for and detecting quality issues [25]. Since nursing home staff rather than an independent authority assess residents' health status, risk of detection bias exists. Thirdly, variation in clinical quality indicators may be due not only to changes in quality, but also in risk or error [26]. To cope with this issue, different risk adjustment techniques are used. While previous studies of nursing home quality mainly use adjustment methods at the facility level [27–29], more recently, risk adjustment is performed at the individual level when data are available. Different approaches include stratification, covariate models [30], and standardization [31]. For some clinical indicators of quality that are considered particularly relevant in detecting the presence of problematic cases of quality shortcomings, no risk adjustment is required. Among these are the presence of daily physical restraints [12], dehydration, and fecal impaction [26, 32]. The main issue of risk adjustment techniques is that they may only partially capture residents' risk factors, resulting in biased estimates of quality coefficients. Risk adjustment is also of concern when risk adjustment factors are themselves a function of quality. In these cases, quality scores could be over-adjusted, giving credit for poor quality [33].

Empirical evidence on the impact of quality on costs

The literature on nursing home costs is extensive, but only marginally addresses quality of care. One challenge lies in the measurement of quality. Due to the absence of clinical indicators of quality, studies mainly use non-clinical measures, such as the number of deficiency citations, information about staffing (e.g., turnover rate or skill characteristics), or mortality rates. Others rely on modeling quality as a latent variable [34, 35].

Empirical models using non-clinical quality measures mainly focus on the impact of specific factors on costs, such as market structure, forms of organization, or reforms implemented in the nursing home sector. Quality measures are usually introduced as control factors. Some of these studies use staffing information [18, 21, 36] or

Table 1 Classification of quality indicators according to the SPO framework developed by Donabedian [9]

Structure	Process	Outcome	Consumer-reported indicators
Room size	Staffing information	Mortality rates	Resident satisfaction
Equipment	Mistakes rate	Hospitalization	Family satisfaction
Staffing levels	RAI quality indicators	RAI quality indicators	Deficiency citations

Table 2 Overview of selected studies investigating the relationship between costs and quality in nursing homes

Study	Quality indicators	Empirical strategy (data)	Effects on costs
Mukamel and Spector [47]	INC of functional decline	Weighted least squares model (cross-sectional data)	Inverted U-shaped
	INC of bedsores		Inverted U-shaped
	Mortality		Inverted U-shaped
Laine et al. [48]	PR of pressure ulcers	Stochastic frontier models (cross-sectional data)	Positive effects
	PR of weekly use of depressants		No significant effects
	PR of depression w/o treatment		No significant effects
Laine et al. [49]	PR of depression w/o treatment	Stochastic frontier models (cross-sectional data)	No significant effects
	PR of pressure ulcer		Positive effects
	PR of use of depressants		Positive effects
	PR of use of physical restraints		No significant effects
Weech-Maldonado et al. [50]	Worsening of pressure ulcers	Two-stage least squares models (cross-sectional data)	Inverted U-shaped
	Mood decline		U-shaped
Wodchis et al. [3]	PR of physical restraints	Random and fixed effects models (panel data)	Negative effects
	PR of use of depressant		No significant effects
	PR of urinary incontinence		Positive effects
	INC of urinary incontinence		U-shaped
	PR of skin ulcers		No significant effects
	INC of skin ulcers		Negative effects
	PR of pain		No significant effects

INC incidence, *PR* prevalence

deficiency rates [37]. Another strand of literature exploits determinants of quality variability. Factors considered include the impact of state regulations [38, 39], ownership form [40, 41], competition [42–45], and financial performance [46].

We focus our review on studies that use clinical indicators derived from the RAI to investigate the relationship between costs and quality. The main contribution of these studies is summarized in Table 2, where details on the choice of quality indicators, the empirical approach, and the results are presented.

Mukamel and Spector [47] analyze nursing homes in New York State using regression-based risk adjustment. The authors report an inverted U-shaped relationship between costs and quality. An important contribution to the cost-quality relationship is provided by Laine et al. [48, 49], who implement stochastic frontier models for the Finnish long-term care sector. The prevalence of pressure ulcers is the only quality indicator positively associated with technical inefficiency. Laine et al. [49] provide a similar cross-sectional analysis that shifts the focus from productive efficiency to cost efficiency. The mean values of the indicators over a three-year period are taken without risk adjustment. The results show that a worse outcome in terms of higher prevalence of pressure ulcers is associated with higher costs, while poor process quality measured by the weekly use of antidepressants and hypnotics is associated with higher inefficiency. However, the impact of these quality indicators is relatively low.

Weech-Maldonado et al. [50] investigate the impact of quality on costs in US nursing homes. Using cross-sectional data from around 750 facilities, they test the inverted U-shaped theory. Indicators are adjusted for risk using the covariates model [51]. To our knowledge, this is the only study that addresses endogeneity by instrumenting the quality indicators with county-level variables associated with nursing home demand (e.g., poverty rate over 65, female older than 75 years, education levels, mortality rate,

Medicare inpatient days). However, the validity of the instruments is not tested. The results show an inverted U-shaped relationship between costs and pressure ulcers. An opposite pattern arises for mood decline, showing that different indicators of quality may lead to different types of relationships. Additional evidence based on data from Ontario, Canada, is provided by Wodchis et al. [3], who estimate panel data models. The analysis shows a negative relationship between costs and daily use of physical restraints, as well as worsening incontinence. Antipsychotic use, the prevalence of pressure ulcers, and the prevalence of severe pain are not statistically significant.

Most of the studies presented above find correlation between some quality indicators and costs. However, the association is weak and the approaches used are hardly comparable. The majority of these studies use a cross-sectional design and do not account for unobserved heterogeneity that may affect both costs and quality. Unobserved heterogeneity may represent a serious problem in analyses of costs-quality relationship due to the difficulty in measuring quality. If the risk adjustment technique used in cross-sectional studies does not capture the facility-specific features perfectly, then the results may be biased. Also, only a few studies address the potential endogeneity of quality, and virtually no test is provided on the validity of the instruments.

In the following section, we propose an empirical approach to investigate the relationship between costs and quality using data from Swiss nursing homes. The main novelty of this approach is the inclusion of process and outcome quality measured by composite or single quality indicators into a cost function. As compared to previous studies, we are also able to control for omitted variable bias by exploiting the panel structure of our data.

Model specification and data

Choice of quality indicators

Quality indicators measure adverse events such as the use of antipsychotic drugs, injuries, bedridden residents, and pressure ulcers. To select appropriate quality indicators from the 22 available in our data set, we consider two approaches. The first approach combines quality indicators to obtain composite measures of process and outcome quality. Conversely, the second approach selects single quality indicators of process and outcome.

Combining different quality indicators, as suggested by organizations including the US Institute of Medicine [52], allows us to condense the multidimensional nature of quality, limit the number of variables included in an econometric model, and overcome possible arbitrariness in the choice of quality indicators. However, combining different quality indicators requires a weighting mechanism, which may be subject to criticism. Differences in the number of eligible residents for different quality events across facilities may represent a serious problem in obtaining a composite measure of quality. To overcome this problem, quality indicators can be adjusted before aggregation to increase comparability across facilities. The avoidable number of residents potentially exposed to different quality events may offer a valid solution for adjustment. As an alternative, one can generate composite indicators using a principal component analysis (PCA), where many single indicators are reduced to a small number of orthogonal components (see for details [53, 54]). Since a composite measure of quality makes it difficult to identify the factors affecting costs, we rely on Donabedian's classification of quality and derive separate composite indicators for both process and outcome. This allows us to identify the effect of the two quality dimensions separately.

To derive composite indicators of process and outcome quality, we then use two methods. First, we weight each quality indicator by the number of residents exposed to a given quality event within each nursing home. The second method applies PCA to single quality indicators of process and outcome to obtain a few orthogonal components, which can be used as composite measures of quality.[3] PCA is a statistical procedure that converts the observations of possibly correlated single quality indicators into a set of linearly uncorrelated variables, called principal components, through an orthogonal transformation. Each succeeding component accounts for as much of the variability in the data as possible under the constraint that it is uncorrelated with the previous components. As a rule of thumb, components with eigenvalues higher than 1 are generally considered. In our case, we consider the first two components to approximate composite measures of process and outcome quality. We have also replicated our analysis using the third component of outcome quality since its eigenvalue is slightly higher than 1, but the results are unchanged.

As stated above, the second approach to select appropriate quality indicators is based on single quality indicators. Single quality measures are probably more reliable and meaningful than composite measures. However, a selection process is needed to limit the number of indicators used in an econometric model. Our selection process

[3] Most of our quality indicators include 173 observations. However, for a few of them, information was collected for only two years. To maximize the number of observations used in the following econometric analysis, we dropped four single quality indicators with missing values (see Table 4 for details).

builds on two strands of literature: the medical recommendations literature and the medical-statistical literature.

Regarding medical recommendations, we consult the numerous lists of recommended indicators used in benchmarking analyses of nursing homes [23, 25]. From the medical-statistical literature, we derive three main criteria that should be satisfied for the empirical analysis (see for instance [49]): a relatively large variation in the quality scores, the absence of multicollinearity between the indicators and other variables, and a relatively large number of observations from which the quality indicators are calculated. The latter criterion is motivated by statistical properties, since some quality indicators capture the onset of rare events. In these cases, the relevant question is whether the observed frequency of the event can be considered a "true score", or whether it is driven by random shocks. Indeed, standard errors of rare events are large and generate problems in the comparison of quality among facilities. Generally, the minimum number of observations for benchmarking is 20 [12].

Based on these criteria, we select two process quality indicators and two outcome quality indicators. The two indicators of process are the presence of antipsychotic use for low-risk residents and the daily use of physical restraints. The two indicators of outcome include the prevalence of weight loss and the prevalence of severe pain. We also control for time-invariant quality features in the structure of nursing homes through the econometric specification of our model (see "Econometric estimation and results").

Detailing the cost function

In order to identify the impact of quality on costs, we consider a cost model that includes quality indicators as derived in "Choice of quality indicators". Total costs are a function of output (Y), measured by the number of patient-days of nursing home care, prices for labor, capital, and material (P_l, P_k, P_m), the institutional form of the nursing home (IF), the case-mix of residents (MIX), the nursing staff ratio (SR), a vector of process and outcome quality indicators ($\mathbf{q}$), and a time trend (τ) that captures technological progress:[4]

$$C = f(Y, P_l, P_k, P_m, \mathrm{IF}, \mathrm{MIX}, \mathrm{SR}, \mathbf{q}, \tau). \quad (1)$$

The price of labor is calculated as the weighted average wage of different professional categories employed in the nursing home (doctors, nurses, administrative and technical staff). The price of capital is calculated as the sum of mortgage costs, amortization, and costs related to capital purchases divided by the capital stock, which is approximated by the number of beds. The price for material and meals is computed by taking the remaining costs and dividing them by the number of meals provided each year. This item mainly includes costs for food, energy, and administrative expenses.

The main difference between nonprofit nursing homes is in their institutional form. Public-law nursing homes are public administrative units without a separate judicial status from the local public administration. Conversely, private-law nursing homes usually take the form of a foundation. We include a dummy for the institutional form (IF) equal to 1 when the nursing home is a public-law organization, and 0 otherwise.

SR is the nursing staff ratio, i.e., the ratio between the number of nurses employed in a nursing home and the number of nurses that should be employed according to the guidelines of the regulator (prescribed amount of staff). Because nursing care is a labor-intensive service, staffing levels have been recognized as a good indicator for (structure) quality [24]. Note, however, that our indicator is conceptually different from other quality indicators related to staffing levels, since it captures deviations from the prescribed number of nurses.

The vector of process and outcome quality indicators ($\mathbf{q}$) leads to three different model specifications. In Model 1, the quality vector includes four composite indicators, two for process quality ($Q^{\mathrm{pc1}}_{\mathrm{process}}$ and $Q^{\mathrm{pc2}}_{\mathrm{process}}$) and two for outcome quality ($Q^{\mathrm{pc1}}_{\mathrm{outcome}}$ and $Q^{\mathrm{pc2}}_{\mathrm{outcome}}$), derived from PCA. As explained above, these are the two succeeding principal components of all the observed quality scores with the highest eigenvalues. In Model 2, the quality vector is represented by two composite indicators (Q_{process} and Q_{outcome}) derived using weights according to the number of residents exposed to different quality inputs. Finally, Model 3 includes a vector of four single quality measures: two process quality indicators—the prevalence of antipsychotic use for low-risk residents (Q_{antips}) and daily use of physical restraints use (Q_{restr})—and two outcome quality indicators—weight loss (Q_{weight}) and severe pain (Q_{pain}).

Q_{antips} is risk-adjusted based on the stratification approach, whereas Q_{restr} is a sentinel indicator, and as such, no risk adjustment is required [12]. Due to lack of data at the resident level, we further control for case-mix differences using an index at the facility level (MIX). This is a cardinal index that measures the average patient's need in terms of daily hours of personal and medical care, and is calculated on yearly basis by the regulator. Patients are classified in one out of five categories according to their severity level. A value between 0 and 4 is assigned, where

[4] In a non-competitive environment such as the Swiss one, there is no reason to assume that nursing homes minimize costs. In this case, the estimated costs function is a "behavioral cost function" [55] and can still be used to make a comparison among firms.

higher values indicate more severe cases.[5] We expect this case-mix indicator to be correlated with patients' risk factors that are not observable. Moreover, any unobserved facility-specific risk factor features are captured by the individual effects. We acknowledge that the risk-adjustment system used in this analysis may be less precise than adjustments based on clinical information at the individual level. However, as previously discussed, even complex systems of risk adjustment present serious shortcomings.

For the estimation of the cost model in Eq. (1), we use a log-log functional form. When choosing the functional form, parsimony in the number of coefficients to be estimated is traded off against flexibility. A translog functional form would require interacting all quality indicators with the production factors, leading to an important loss of degrees of freedom.[6,7]

Input prices and total costs are divided by the material price in order to satisfy the homogeneity condition in input prices.[8] The log-log form of Eq. (1) is:

$$\ln\left(\frac{C}{P_{\mathrm{m}}}\right) = \delta_0 + \delta_Y \ln Y + \delta_{P_{\mathrm{l}}} \ln\frac{P_{\mathrm{l}}}{P_{\mathrm{m}}} + \delta_{P_{\mathrm{k}}} \ln\frac{P_{\mathrm{k}}}{P_{\mathrm{m}}} + \delta_{\mathrm{IF}}\mathrm{IF} + \delta_{\mathrm{MIX}} \ln \mathrm{MIX} + \delta_{\mathrm{SR}}\mathrm{SR} + \delta_q \mathbf{q} + \delta_t \tau + \varepsilon, \quad (2)$$

where δ_q is the vector of quality parameters, ε is the error term that contains individual effects δ_i. The individual subscript i and the time subscript t are omitted for simplicity.

The estimation of the cost function in Eq. (2) is based on the assumption that output, input prices, and quality are exogenous variables. In the case of nursing homes included in the sample, output is likely to be exogenous because nursing homes have to accept all residents in a given residential area, and residents do not have free choice of the facility. Also, the excess of demand due to subsidized prices leads to occupation rates of about 100 %. For the same reasons, the case-mix is also likely to be exogenous. Moreover, the reimbursement system is linked to the nursing home-specific case-mix, which limits incentives to attract less costly patients. Input prices can be considered exogenous because nursing homes have to follow the guidelines imposed by the regulator.

As with respect to quality, it is important to distinguish between the nursing staff ratio and clinical quality indicators derived from the RAI. The nursing staff ratio is strongly regulated by the canton, and nursing homes are not allowed to deviate significantly from the optimal staff size. Therefore, we can exclude the presence of endogeneity.[9] The potential endogeneity issue of unregulated clinical indicators will be discussed later, in "Econometric estimation and results".

Data and descriptive statistics

We merge two data sets on costs and quality of nursing home residents in southern Switzerland (Canton Ticino), which were provided by the regulator. The first data set includes yearly use of resources at the organization level extracted from the annual reports of nursing homes. It includes 45 nursing homes over a 10-year period, from 2001 to 2010. The second data set contains information derived from the MDS on 22 clinical quality indicators at the organization level for the period 2006–2010, excluding the year 2008.[10] Due to missing values in the data set, no quality scores are available for three nursing homes for the years 2006 and 2007. The total number of observations is 173 for the models with composite quality indicators. For the model with single quality indicators, we exclude observations for which the denominator of the quality scores is less than 20. This leads to a loss of ten observations.

In Tables 3 and 4, we provide descriptive statistics for the main costs and quality variables. The data show that on average, a resident day costs 247 Swiss francs (SFr.). The difference between the minimum and the maximum cost is almost SFr. 200. This may be due to differences in the output, as the number of resident days ranges between almost 9000 and more than 64,000. The average resident case-mix is 3.1, with important differences among nursing homes (0.80–3.83). The average price of labor is approximately SFr. 81,000, and nursing homes are highly homogenous in this respect. The prices of capital and

[5] Note that this is not the Resource Utilization Group (RUG)'s classification system of residents. As compared to the RUG system, our case-mix measure is not derived from the MDS. The main advantage is that case-mix differences are less likely to reflect quality levels.

[6] In a preliminary analysis, we also estimated: (1) a full translog cost model and (2) a hybrid translog cost model. In the hybrid translog cost function, quality indicators were included only in linear form. The results of the full translog were not satisfactory, probably due to multicollinearity problems and the loss of degrees of freedom. The results of the hybrid cost function were very similar to those obtained with the log-log functional form.

[7] Squared terms for quality indicators were also considered in a separate analysis to test the presence of a non-linear relationship between quality and costs. The results did not show evidence of non-linear relationship.

[8] The cost function is linear homogenous of degree 1 in input prices when a 10 % increase in all input prices leads to a 10 % increase in total cost.

[9] The Durbin–Wu–Hausman test performed using the lagged SR as instrumental variable does not reject exogeneity at the 99 % level.

[10] Four of these indicators are risk-adjusted based on the stratification approach. This means that they are calculated separately for high-risk and low-risk patients. In these cases, we use the low-risk indicators.

Table 3 Descriptive statistics of costs, output, and structure variables

Variable	Description	Obs	Mean	SD	Min	Max
C	Average cost (SFr./Y)	440	246.90	25.66	179.48	377.52
Y	Annual resident days	519	25,434	10,231	8955	64,275
P_{l}	Average labor price in SFr. per employee per year	519	80,817	5196	63,363	97,512
P_{k}	Average capital price in SFr. per bed	440	5735	2528	1054	22,981
P_{m}	Average material price in SFr. per meal	440	9.23	4.74	5.16	103.25
MIX	Average dependency index	519	3.10	0.34	0.80	3.83
SR	Nursing staff ratio	519	0.96	0.09	0.74	1.55

All monetary values are in 2005 Swiss francs (SFr.), adjusted by the national Consumer Price Index

Table 4 Descriptive statistics of process and outcome quality indicators

Quality indicator	Description	Obs	Mean	SD	Min	Max
Process						
$Q_{\mathrm{depressionNT}}$	PR of depression symptoms without treatment	173	31.45	12.54	3.6	60.0
Q_{drugs}	Prevalence of use of 9 or more medications	173	41.46	14.30	0.0	75.0
$Q_{\mathrm{catheters}}$	INC of indwelling catheters	173	5.38	5.25	0.0	27.9
Q_{gavage}	PR of enteral feeding	130	1.46	2.67	0.0	17.5
$Q_{\mathrm{antipsyHR}}$	PR of antipsychotic use: HR	173	51.47	23.22	0.0	100.0
$Q_{\mathrm{antipsyLR}}$	PR of antipsychotic use: LR	173	31.67	12.45	7.7	87.5
$Q_{\mathrm{restraints}}$	PR of daily physical restraints	173	19.56	9.76	0.0	50.0
Q_{activity}	PR of little or no activity	173	52.73	19.74	6.7	100.0
Q_{process}	Composite indicator of process quality	173	25.83	5.01	12.9	40.3
$Q_{\mathrm{process}}^{\mathrm{pc1}}$	First principal component of process quality	173	0.00	1.23	−3.2	3.3
$Q_{\mathrm{process}}^{\mathrm{pc2}}$	Second principal component of process quality	173	0.00	1.16	−3.6	3.2
Outcome						
Q_{injuries}	INC of injuries	159	2.39	6.64	0.0	50.0
Q_{falls}	PR of falls	173	9.20	5.70	0.0	24.4
$Q_{\mathrm{behaviourHR}}$	PR behavior problems: HR	173	34.44	13.80	0.0	78.9
$Q_{\mathrm{behaviourLR}}$	PR behavior problems: LR	173	15.30	13.33	0.0	70.0
$Q_{\mathrm{depression}}$	PR of depression symptoms	173	56.41	16.47	14.8	100.0
$Q_{\mathrm{incontinenceHR}}$	PR of bowel incontinence: HR	173	80.23	19.86	0.0	100.0
$Q_{\mathrm{incontinenceLR}}$	PR of bowel incontinence: LR	173	30.54	19.40	0.0	100.0
$Q_{\mathrm{incontinence}}$	PR of bladder incontinence without a toileting plan	173	59.88	25.91	0.0	100.0
Q_{weight}	PR of weight loss	173	6.75	5.10	0.0	27.3
Q_{bedfast}	PR of bedfast residents	172	8.48	6.75	0.0	29.6
Q_{los}	INC of decline in late-loss activities of daily living	154	30.79	26.12	0.0	100.0
Q_{ulcersHR}	PR of pressure ulcers: HR	173	11.67	9.24	0.0	50.0
Q_{ulcersLR}	PR of pressure ulcers: LR	173	2.90	5.01	0.0	28.6
Q_{pain}	PR of important pain	173	21.24	11.99	0.0	61.1
Q_{outcome}	Composite indicator of outcome quality	173	17.68	4.08	6.8	28.6
$Q_{\mathrm{outcome}}^{\mathrm{pc1}}$	First principal component of outcome quality	173	0.00	1.31	−3.4	5.0
$Q_{\mathrm{outcome}}^{\mathrm{pc2}}$	Second principal component of outcome quality	173	0.00	1.17	−2.7	3.0

INC incidence, *PR* prevalence, *HR* high-risk, *LR* low-risk

material show higher variation, from SFr. 1054 to almost SFr. 23,000 and from SFr. 5.16 to around SFr. 103, respectively. These differences are due to renovation or enlargement investments. At the approximation point, the shares of capital, material, and labor costs are 6.5, 12.1, and 81.4 %, respectively.

Regarding quality indicators, the data show that the nursing staff ratio is very close to 1, as expected. Variations larger than 10 % are possible only for very short periods. On average, 32 % of low-risk patients use antipsychotics, but in some nursing homes this value reaches 88 %, suggesting that serious problems may exist within the production process of nursing home care. The average prevalence of daily use of physical restraints is around 20 % and ranges between 0 and 50 %. Regarding outcome quality, the average prevalence of residents who lost weight unexpectedly is about 7 %, and this percentage ranges between 0 and 27 %. Finally, the prevalence of residents suffering from severe pain is 21 % on average, but reaches more than 60 % in some cases.

An interesting question is whether quality domains are correlated. This may affect the selection process of appropriate composite quality scores as well as the choice of single quality indicators to be included in the econometric analysis. We compute the correlation among indicators (including the staff ratio) and Kendall's rank correlation coefficient [56]. The latter measures the similarity of the ordering of nursing homes when these are ranked according to quality scores. Both measures indicate a low correlation between quality indicators (<25 %). This could potentially undermine the use of composite quality scores derived from PCA (Model 1). However, our strategy of measuring quality using three different approaches comes out stronger. Indeed, the use of other composite quality scores (Model 2) not derived from PCA and the use of single quality indicators (Model 3) may offer an answer to this criticism. Meanwhile, the use of composite quality indicators and of a small number of single quality indicators ensures that collinearity between quality scores is not an issue.

Econometric estimation and results

Panel data models

When analyzing the impact of process and outcome quality on costs, two main issues that are likely to bias the results may arise: omitted variables and simultaneity of quality and costs. We exclude the ordinary least squares (OLS) estimator due to the presence of unobserved heterogeneity (shown by both F test and Breusch–Pagan test) and use panel data models with individual effects.

The results of the estimation of the three models with combined and single quality indicators described in "Detailing the cost function" are reported in Table 5. In all models, standard errors are corrected using the cluster-robust estimator based on Stock and Watson [57] and Kezdi [58].[11] Both the FE and the RE models have potential advantages and disadvantages, and the model choice involves a tradeoff between bias and variance [60]. Both approaches address the omitted variables issue. The RE model treats the individual effects as stochastic parameters, therefore assuming independence with the other covariates. When this assumption does not hold true, the RE estimates are biased. Instead, the fixed effects model treats the individual effects as fixed parameters and allows the individual effects to be partially correlated with regressors, thus accommodating a limited form of endogeneity deriving from constant omitted variables [61]. This feature is particularly appealing in studies of costs and quality due to unmeasurable dimensions that are likely to affect the relationship. The Hausman test casts doubts on the RE estimates, since it rejects at the 5 % level the hypothesis that the individual-specific error terms are uncorrelated with the explanatory variables, i.e., the RE estimator may be inconsistent (see Cameron and Trivedi [62] for details). As compared to the RE model, the FE model could suffer from lack of robustness in the case of small sample size or small within variation. However, our estimates show that the coefficients of interest (quality indicators) are very stable independently of the model specification. As expected, the variance is smaller in the RE estimates. Given that the percentage of within variation of the variables of interest with respect to the overall variation is satisfactory, the fixed effects estimates should be fairly precise [62].

The small sample size may explain the difference in the magnitude of some coefficients (Y, MIX, and SR) that increase slightly in the RE estimates. However, the sign and statistical significance of all coefficients are basically unchanged, suggesting that FE estimates are unbiased. The only exceptions are a measure of process quality (Q_{restr}), which becomes significant at the 10 % level in the RE specification of Model 2, and a measure of outcome quality ($Q^{\text{pc2}}_{\text{outcome}}$), which becomes more significant in the FE specification of Model 1. Generally, the similarity of the random effects and the fixed effects estimates suggests a low correlation between the individual effects and our covariates.

Note that the estimated parameters are very similar across the three models. Consider first the main variables of interest: the quality indicators. The nursing staff ratio (SR) is highly significant. As expected, the higher the relative number of nurses working in a nursing home, the higher the costs. The estimated coefficient is stable across the three

[11] Kezdi [58] states that a sample of 50 clusters is close enough to infinity for accurate inference if the number of observations per cluster is not too small. A cluster is considered small if it contains less than five observations [59]. In our case, the significance of the coefficients remains unchanged when standard errors are clustered.

Table 5 Estimation results of fixed-effects (FE) and random-effects (RE) cost models with composite and single quality indicators

	Model 1				Model 2				Model 3			
	Composite quality indicators using PCA				Composite quality indicators using no. of exposed				Single quality indicators			
	FE		RE		FE		RE		FE		RE	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Constant	−3.369***	0.475	−4.570***	0.316	−3.477***	0.505	−4.628***	0.321	−3.621***	0.541	−4.657***	0.328
Y	0.728***	0.040	0.845***	0.020	0.736***	0.041	0.847***	0.019	0.751***	0.046	0.853***	0.019
P_{l}	0.911***	0.024	0.905***	0.026	0.911***	0.025	0.904***	0.027	0.916***	0.026	0.910***	0.027
P_{k}	0.066***	0.014	0.067***	0.013	0.067***	0.015	0.068***	0.014	0.059***	0.015	0.059***	0.013
MIX	0.199***	0.060	0.244***	0.054	0.202***	0.064	0.247***	0.055	0.219***	0.079	0.254***	0.060
SR	0.493***	0.061	0.530***	0.061	0.482***	0.066	0.516***	0.063	0.480***	0.071	0.513***	0.069
Public	–	–	−0.005	0.016	–	–	−0.005	0.015	–	–	−0.007	0.015
$Q^{\mathrm{pc1}}_{\mathrm{process}}$	−0.001	0.003	−0.001	0.003	–	–	–	–	–	–	–	–
$Q^{\mathrm{pc2}}_{\mathrm{process}}$	−0.006	0.005	–0.003	0.004	–	–	–	–	–	–	–	–
$Q^{\mathrm{pc1}}_{\mathrm{outcome}}$	0.007**	0.003	0.007**	0.003	–	–	–	–	–	–	–	–
$Q^{\mathrm{pc2}}_{\mathrm{outcome}}$	0.006**	0.003	0.005*	0.003	–	–	–	–	–	–	–	–
Q_{process}	–	–	–	–	−0.048	0.070	0.026	0.066	–	–	–	–
Q_{outcome}	–	–	–	–	0.240**	0.094	0.212**	0.086	–	–	–	–
Q_{antips}	–	–	–	–	–	–	–	–	0.026	0.025	0.034	0.025
Q_{restr}	–	–	–	–	–	–	–	–	−0.064	0.042	−0.071*	0.040
Q_{pain}	–	–	–	–	–	–	–	–	0.056**	0.027	0.061**	0.028
Q_{weight}	–	–	–	–	–	–	–	–	0.102**	0.043	0.098**	0.044
t	0.003	0.002	0.003	0.002	0.006**	0.002	0.005**	0.002	0.005***	0.002	0.004***	0.002
R^2 (Adj.)	0.978		0.982		0.977		0.982		0.980		0.983	
N	173		173		173		173		163		163	

Significance levels: * 10 %, ** 5 %, *** 1 %

models. In Model 1 and Model 2, we consider composite quality indicators. Note that outcome quality ($Q^{pc1}_{outcome}$, $Q^{pc2}_{outcome}$, $Q_{outcome}$) exhibits a negative (positive sign) and significant effect on costs in both models, although the magnitude of the effect is stronger when composite quality indicators are derived using weights according to the number of residents exposed to different quality aspects (Model 2). Conversely, process quality shows an opposite (negative sign) effect on costs, although the impact is not significant. These results are in accordance with those obtained with single quality measures (Model 3). We observe a negative and significant association between costs and outcome quality measured by the prevalence of weight loss (Q_{weight}) and the prevalence of severe pain (Q_{pain}).[12] This means that worsening outcome measures lead to increased costs, while better control of patients' outcomes reduces nursing home costs. Instead, process quality measured by the daily use of physical restraints (Q_{restr}) and the prevalence of antipsychotic use for low-risk residents (Q_{antips}) do not seem to have a significant impact on costs.

Note also that the other coefficients are very similar across the three models. The coefficient of output (Y) measures the total costs elasticity with respect to output. A value lower than 1 suggests the presence of unexploited economies of scale. In our case, an increase in output by 10 % in the number of patient-days increases total costs by roughly 7–8 %. As expected, more severe patients (MIX) are more costly to treat. The coefficient can also be interpreted as a cost elasticity. An increase in the level of patients' severity by 10 % significantly increases costs by around 2 %. The above findings on the effect of outcome quality on costs may be questioned if less costly NHs select patients in better health status, resulting in better outcome indicators for quality. This supposes that the case-mix variable (MIX) does not fully capture the information relative to patients' health status. However, as explained in the "Introduction" section, patient selection is a negligible factor in our setting, since individuals are assigned to the NH of the former place of residence. Still, as a robustness check, we grouped observations in five categories of case-mix (20 percentiles) and did not find any evidence of a systematic increase/decrease in quality when moving from less severe to more severe residents. Finally, we also instrumented the case-mix with the spatial lag and the proportion of elderly population in the community area. The instrumental variables (IV) approach could not reject the hypothesis of exogenous case-mix in all our models (estimation results and endogeneity tests are available upon request).

The cost function is monotonically increasing in the vector of input prices, since input price coefficients (P_l and P_k) are positive and significant. Also, these coefficients provide information on the percentage of labor and capital costs over total costs of a representative NH. The share of labor costs (P_l) is estimated between 91 and 92 %, while the estimated share of capital (P_k) is between 6 and 8 % . The institutional form (IF) is dropped in fixed effects regressions because of time invariance, but it is not significant in random effects regressions.[13] The time trend (t) is statistically significant in Models 2 and 3, but the coefficient is very small. Total costs of nursing home care remained relatively constant over the time period considered in the analysis.

The issue of simultaneity may arise from the fact that costs and quality are codetermined. However, the results discussed so far are not expected to be significantly affected by simultaneity bias for two main reasons. First, our panel is relatively short. Second, as explained above, the nursing home sector under analysis is highly regulated. Nevertheless, to increase the robustness of our findings, we discuss the issue of endogenous quality indicators in detail in the next section.

Instrumental variable models

We believe that simultaneity between costs and quality is unlikely. Even in the case of endogeneity, however, the estimation bias due to quality endogeneity would be very limited. This is because of the institutional setting of the nursing home sector and the strong regulation system. Nursing home activities are regulated by the local government in a relatively effective way. Nonetheless, in order to test potential endogeneity, we consider IV approaches using the efficient generalized method of moments (GMM) combined with the fixed effects model. The GMM approach has the advantage of consistency in the case of arbitrary heteroskedasticity and shows higher flexibility than two-stage least squares (2SLS), in particular to test the validity of the instruments. The GMM approach is preferred, since it allows error clustering for panel data and provides a battery of tests to check the validity of the instruments.

A valid instrument must satisfy two requirements: the instrument z must be correlated with the endogenous variable x, $Cov(z,x) \neq 0$, and uncorrelated with the error term u, $Cov(z,u) = 0$. In the case of multiple endogenous regressors, the Shea partial R^2 [63] measure should be used

[12] Note that the correlation between outcome quality indicators in Model 3 is relatively low (0.17). Clearly, the correlation between outcome quality indicators obtained using PCA (Model 1) is zero, because different components are orthogonal.

[13] For comparison purposes, we also ran RE regressions without the institutional form (IF). The size of the coefficients remains unchanged (estimates not reported).

Table 6 Estimation results of second-stage IV-GMM cost models with fixed effects

	Model 1 IV-GMM		Model 2 IV-GMM		Model 3 IV-GMM	
	Composite quality indicators using PCA		Composite quality indicators using no. of exposed		Single quality indicators	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Y	0.776***	0.037	0.715***	0.042	0.850***	0.060
P_l	0.928***	0.017	0.916***	0.025	0.912***	0.017
P_k	0.060***	0.012	0.067***	0.013	0.082***	0.023
MIX	0.235***	0.061	0.182***	0.063	0.305***	0.093
SR	0.510***	0.065	0.488***	0.072	0.491***	0.076
$Q^{pc1}_{process}$	0.013*	0.007	–	–	–	–
$Q^{pc2}_{process}$	−0.013	0.011	–	–	–	–
$Q^{pc1}_{outcome}$	0.010*	0.005	–	–	–	–
$Q^{pc2}_{outcome}$	0.009***	0.003	–	–	–	–
$Q_{process}$	–	–	−0.300	0.240	–	–
$Q_{outcome}$	–	–	0.596*	0.343	–	–
Q_{antips}	–	–	–	–	0.177*	0.092
Q_{restr}	–	–	–	–	0.027	0.103
Q_{pain}	–	–	–	–	0.016	0.121
Q_{weight}	–	–	–	–	0.312	0.268
t	0.001	0.003	0.004**	0.002	0.008***	0.003
$Adj.R^2$	0.978		0.964		0.972	
N	125		173		125	
Overidentifying restrictions test[a]						
χ^2	8.656		0.744		1.932	
p value	0.194		0.863		0.926	
Endogeneity test[b]						
χ^2	1.204		1.448		5.817	
p value	0.878		0.485		0.213	

Excluded instruments: average population density and distance, proportion of young, adults and elderly people in the neighbouring area (Models 1, 2, and 3), global and local spatial lags of quality indicators (Models 1 and 3)

Significance levels: * 10 %, ** 5 %, *** 1 %

[a] For all instruments jointly. Hansen J statistic $\sim\chi^2(6)$ for Model 1 and 3, $\chi^2(3)$ for Model 2

[b] Regressors tested: $Q^{pc1}_{process}$, $Q^{pc2}_{process}$, $Q^{pc1}_{outcome}$, $Q^{pc2}_{outcome}$ in Model 1; $Q_{process}$ and $Q_{outcome}$ in Model 2; Q_{antips}, Q_{restr}, Q_{pain} and Q_{weight} in Model 3

to test the first condition, as this takes into account the intercorrelation among the instruments.[14] However, this does not exclude the possibility of weak instruments. The second condition can be tested when there are more instruments for an endogenous variable. In this case, the C-statistic, also called "difference-in-Sargan" statistic, can be used [64].

As shown in previous studies [3, 47], good instruments for quality are lacking. Moreover, finding good instruments for several quality indicators is even more challenging. We rely on three hypotheses. First, visits by residents' relatives exert pressure on the management staff of the nursing home to keep adequate levels of quality. Hence, we identify two variables: the weighted average distance (travel time) between residents' location and the nursing home facility, and the weighted population density of the area served by the nursing home. The relative distance variable has previously been used by other authors in the nursing home literature and is considered a valid exclusion restriction [40, 65, 66]. The second hypothesis assumes that the quality offered by the nursing home depends on the average quality offered by surrounding nursing homes. We build a variable to capture pressure from other nursing homes located in geographical proximity. For each year

[14] The F diagnostic for weak instruments for the joint significance of the instruments in first-stage regressions does not recognize situations in which some instruments are good while others are weak.

and nursing home, pressure is measured as the average score of quality indicators of nursing homes located in neighbouring districts.[15] Our third hypothesis is that the elderly population living in the area around the nursing home exercises an indirect pressure on quality of home care provided. We then consider the percentages of young, adult, and elderly people in the catchment area of each nursing home. Finally, we also consider lagged values of quality indicators as natural instruments.[16]

The results of IV-GMM estimations are reported in Table 6. The table shows the three IV-GMM models with fixed effects. The results of FE estimations without IV are partially confirmed. Findings are mixed. Process quality indicators become significant at 10 % in a couple of cases, whereas outcome quality indicators lose significance in some cases. The Hausman test suggests no evidence of endogeneity of quality indicators. The Hansen J test indicates that the over-identifying restrictions are valid. *F* tests of excluded instruments in the first stage are passed for most regressors. However, the Shea partial R^2 statistics show that the percentage of variability in quality indicators explained by the instruments is relatively low. Because of the small sample and several potentially endogenous regressors, *F* statistics are not high and the instruments do not appear to be strong enough to safely conclude that quality endogeneity can be excluded.[17] As stated above, addressing endogeneity using multiple quality indicators and many instruments may not be very efficient. Consequently, we also tested exactly identified models with only one quality indicator and one instrument. In these cases, the null hypothesis that the excluded instruments are exogenous cannot be rejected, and the results appear more robust to weak identification.

Conclusions

In the nursing home sector, poor quality represents a main concern, and ongoing discussions are taking place to address this issue. How to increase quality in a context of financial pressure remains an open question. In this paper, we contributed to this debate by investigating the relationship between costs and quality in accordance with the SPO framework developed by Donabedian. We used recently published data on quality indicators derived from the resident assessment instrument and costs of Swiss nursing homes. In addition to structure quality indicators (e.g., nursing staff ratio), we considered single and composite clinical measures of process and outcome quality.

As compared to previous studies, we improved the estimation approach by using panel data models, in particular the fixed effects model to address endogeneity arising from omitted variables. In addition, we instrumented the quality indicators and tackled bias coming from potential simultaneity between costs and quality. While we did not find evidence of simultaneity bias and were able to control for constant omitted variables bias, we could not exclude bias from omitted, time-varying variables.

Our analysis showed evidence of a negative and significant relationship between clinical indicators of outcome quality (e.g., the prevalence of severe pain and the prevalence of weight loss) and total costs. Conversely, we did not find an impact of process quality on costs. Prevalence of daily physical restraint use, as well as the use of antipsychotics, were not found to be statistically significant. Interestingly, process measures typically interpreted as labor-saving and cost-saving factors did not seem to affect costs, while outcome measures did. Finally, structure quality indicators such as staffing levels were strongly associated with higher costs, as shown in previous studies.

A possible explanation for the negative effect of outcome quality on costs is that the use of cost-saving instruments, such as drugs and physical restraints, may initially reduce costs. However, this is only a temporary effect since worsening patient outcomes lead to increased costs of treatment making up more than the initial savings. This explanation may be questioned, since our results are based on a relatively short panel. Note, however, that the large majority of Swiss home care residents (> 60 %) are more than 85 years old and do not spend many years in a nursing home. Therefore, the effects of poor outcome quality on costs are expected to manifest in a relatively short period. Untreated patients may develop more severe dysfunctions, and, consequently, require additional resources in subsequent treatments.

An alternative interpretation relies on the idea of patient selection, i.e., nursing homes that are less costly also select patients in better health states, resulting in better outcome indicators for quality. Although we cannot completely exclude this hypothesis, our regulatory setting and the analysis provide evidence against this interpretation. First, the tight Swiss regulation on individual access to nursing homes makes resident selection highly unlikely. Individuals are assigned to the NH of the former place of residence. Second, we did not find evidence of a systematic trend in

[15] The region considered in the analysis is divided into eight districts: Mendrisio, Lugano, Vallemaggia, Locarno, Bellinzona, Riviera, Blenio, and Leventina. Given that only a few nursing homes are located in northern districts, Vallemaggia, Leventina, and Blenio are pooled together.

[16] Lagged values are an attractive instrument due to the high correlation with the endogenous variable. Nevertheless, caution is necessary in the presence of serial correlation in the data, as this may invalidate the instruments [67]. To test for autocorrelation in the panel data set, we use the test developed by Wooldridge [68, 69].

[17] See, for instance, Hahn et al. [70], for a discussion about weak instruments in the econometric literature.

quality when moving from less severe to more severe residents or of endogenous patient severity in IV regressions.

From a policy point of view, the assessment of the relationship between costs and quality may be valuable in informing payment systems for long-term care. Our results may lead to paradoxical conclusions on the properties of payment systems. Generally, funding schemes for long-term care do not compensate nursing homes for outcome quality. This is the case in the Canton Ticino, Switzerland. At the beginning of the period of analysis (2006), the Canton Ticino introduced a new payment system based on prospective payments (global budget) and started a system of quality measurement. The cantonal authority does not rule out to integrate quality aspects in future revisions of the payment system. Given our results, the current payment system may provide adequate incentives for cost containment if managers are aware of the negative relationship between outcome quality and costs. Since payments are independent of actual costs, managers may have the incentive to better manage and prevent adverse clinical outcomes such as pain and weight loss to avoid increasing costs. Conversely, under a cost reimbursement system, managers may not have the incentive to prevent adverse outcomes, since additional costs to treat residents when adverse events occur would be covered. However, if managers are imperfectly informed on the relationship between costs and quality, the cantonal authority could consider two options: to improve information available to NH managers on the effects of quality on costs, or to incorporate the effects of quality on costs into the payment mechanism, e.g., by rewarding quality improvements or providing negative financial incentives for poor outcome quality. The latter instrument suggests that incorporating quality aspects into retrospective payment schemes would lead to quality improvements.

To conclude, it is not being advocated that a measure of the impact of quality on costs should be used in a mechanical way to introduce financial incentives in payment schemes. Rather, policymakers could use this as an additional instrument to provide a guide to the relative levels of efficiency. However, it should be noted that our results could be sensitive to the assumptions adopted regarding the econometric approach, the model specification, and data limitations. Also, Swiss payment systems for long-term care are quite heterogeneous across cantons. The investigation of the relationship between costs and quality in long-term care in other regulatory settings was beyond the scope of the present study. Therefore, the contribution to the discussion on optimal design of payment schemes in nursing home care is likely to improve in future research.

References

1. Riedel, M., Kraus, M.: The organisation of formal long-term care for the elderly: results from the 21 European country studies in the ANCIEN Project. Social Welfare Policies, ENEPRI Research report (2011)
2. Di Giorgio, L., Filippini, M., Masiero, G.: Implications of global budget payment system on nursing home costs. Health Policy **115**, 237–248 (2014)
3. Wodchis, W.P., Teare, G.F., Anderson, G.M.: Cost and quality evidence from Ontario long term care hospitals. Med. Care **45**, 981–988 (2007)
4. Lohr, K.N. (ed.): Medicare: A Strategy for Quality Assurance, vol. 1. National Academy Press, Washington, DC (1990)
5. A First Class Service—Quality in the New NHS. Department of Health, London (1997)
6. Council of Europe.: Recommendation on Development and Implementation of Quality Improvement Systems (QIS) in Health Care and Explanatory Memorandum (41st Meeting, 24–26 June). Council of Europe, Strasbourg (1997)
7. WHO.: The World Health Report 2000: health systems: improving performance. World Health Organization, Geneva (2000)
8. Legido-Quigley, H., McKee, M., Nolte, E., Glinos, I.A.: Assuring the quality of health care in the European Union. European Observatory on Health Systems and Policies: Observatory Studies Series No. 12 (2008)
9. Donabedian, A.: The quality of care. How can it be assessed? JAMA **260**(12), 1743–1748 (1988)
10. Zimmerman, D.: Development and testing of nursing home quality indicators. Health Care Financ. Rev. **16**(4), 107–127 (1995)
11. Zimmerman, D.: Improving nursing home quality of care through outcomes data: the MDS quality indicators. Int. J. Geriatr. Psychiatry **18**, 250–257 (2003)
12. Berg, K., Mor, V., Morris, J., Murphy, K.M., Moore, T., Harris, Y.: Identification and evaluation of existing nursing homes quality indicators. Health Care Financ. Rev. **23**(4), 19–36 (2002)
13. Li, Y., Cai, X., Ye, Z., Glance, G.G., Harrington, C., Mukamel, D.B.: Satisfaction with Massachusetts nursing home care was generally high during 2005–2009, with some variability across facilities the productive efficiency and clinical quality of institutional long-term care for the elderly. Health Affairs **32**(8), 1416–1425 (2013)
14. Nakrem, S., Vinsnes, A.G., Harkless, G.E., Paulsen, B., Seim, A.: Nursing sensitive quality indicators for nursing home care: International review of literature, policy and practice. Int. J. Nurs. Stud. **46**, 848–857 (2009)
15. Castle, N.G., Ferguson, J.C.: What is nursing home quality and how is it measured? Gerontologist **50**(4), 426–442 (2010)
16. McKay, N.L.: Quality choice in Medicaid markets: the case of nursing homes. Quarter. Rev. Econ. Bus. **29**(2), 27–40 (1989)
17. Farsi, M., Filippini, M., Kuenzle, M.: Unobserved heterogeneity in stochastic cost frontier models: an application to Swiss nursing homes. Appl. Econ. **37**, 2127–2141 (2005)
18. Farsi, M., Filippini, M., Lunati, D.: Economies of scale and efficiency measurement in Switzerland's nursing homes. Swiss J. Econ. Stat. **144**, 359–378 (2008)
19. Castle, N.G., Engberg, J.: Staff turnover and quality of care in nursing homes. Med. Care **43**(6), 616–626 (2005)
20. Castle, N.G., Engberg, J.: The influence of staffing characteristics on quality of care in nursing homes. Health Res. Educ. Trust **42**(5), 1822–1847 (2007)
21. Dormont, B., Martin, C.: Quality of service and cost-efficiency of French nursing homes. 9th European Conference on Health Economics (ECHE), Zurich, July 18–21 (2012)

22. Spilsbury, K., Hewitt, C., Stirk, L.: The relationship between nurse staffing and quality of care in nursing homes: a systematic review. Int. J. Nurs. Stud. **48**, 732–750 (2011)
23. Rantz, M.J., Hicks, L., Grando, V., Petroski, G.F., Madsen, R.W., et al.: Nursing home quality, cost, staffing, and staff mix. Gerontologist **44**(1), 24–38 (2004)
24. Bostick, J.E., Rantz, M.J., Flesner, M.K., Riggs, C.J.: Systematic review of studies of staffing and quality in nursing homes. J. Am. Med. Dir. Assoc. **7**(6), 366–376 (2006)
25. Mor, V., Berg, K., Angelelli, J., Gifford, D., Morris, J., Moore, T.: The quality of quality measurement in US nursing homes. Gerontologist **43**(2), 37–46 (2003)
26. Arling, G., Karon, S.L., Sainfort, F., Zimmerman, D.R., Ross, R.: Risk adjustment of nursing home quality indicators. Gerontologist **37**(6), 757–766 (1997)
27. Nyman, J.A.: Improving the quality of nursing home outcomes. Med. Care **26**, 1158–1171 (1988)
28. Zinn, J.S., Aaronson, W.E., Rosko, M.D.: The use of standardized indicators as quality improvement tools: an application in Pennsylvania nursing homes. Am. J. Med. Qual. **8**, 72–78 (1993a)
29. Zinn, J.S., Aaronson, W.E., Rosko, M.D.: Strategic groups, performance, and strategic response in the nursing home industry. Health Service Res. **29**, 187–205 (1994)
30. Mukamel, D.B.: Risk-adjusted outcome measures and quality of care in nursing homes. Med. Care **35**(4), 367–385 (1997)
31. Zinn, J.S., Aaronson, W.E., Rosko, M.D.: Variations in the outcomes of care provided in Pennsylvania nursing homes: facility and environmental correlates. Med. Care **31**, 475–487 (1993b)
32. Karon, S.L., Sainfort, F., Zimmerman, D.R.: Stability of nursing home quality indicators over time. Med. Care **37**(6), 570–579 (1999)
33. Mukamel, D.B., Glance, L.G., Li, Y., Weimer, D.L., Spector, W.D., Zinn, J.S., Mosqueda, L.: Does risk adjustment of the CMS quality measures for nursing homes matter? Med. Care **46**(5), 532–541 (2008)
34. Gertler, P.J., Waldman, D.M.: Quality-adjusted cost functions and policy evaluation in the nursing home industry. J. Polit. Econ. **100**, 1232–1256 (1992)
35. Carey, K.: A panel data design for estimation of hospital cost functions. Rev. Econ. Stat. **79**(3), 443–453 (1997)
36. Konetzka, R.T., Yi, D., Norton, E.C., Kilpatrick, K.E.: Effects of Medicare payment changes on nursing home staffing and deficiencies. Health Services Res. **39**, 463–488 (2004)
37. Harrington, C., Woolhandler, S., Mullan, J., Carrillo, H., Himmelstein, D.U.: Does investor ownership of nursing homes compromise the quality of care? Am. J. Public Health **91**(9), 1452–1455 (2001)
38. Bowblis, J.R., Crystal, S., Intrator, O., Lucas, J.A.: Response to regulatory stringency: the case of antipsychotic medication use in nursing homes. Health Econ. **21**, 977–993 (2012)
39. Bowblis, J.R., Lucas, J.A.: The impact of state regulations on nursing home care practices. J. Regul. Econ. **42**, 52–72 (2012)
40. Grabowski, D.C., Feng, Z., Hirth, R.A., Rahman, M., Mor, V.: Effect of nursing home ownership on the quality of post-acute care: an instrumental variables approach. J. Health Econ. **32**, 12–21 (2013)
41. Spector, W.D., Selden, T.M., Cohen, J.W.: The impact of ownership type on nursing home outcomes. Health Econ. **7**, 639–653 (1998)
42. Castle, N.G., Liu, D., Engberg, J.: The association of nursing home compare quality measures with market competition and occupancy rates. J. Healthcare Qual. **30**(2), 4–14 (2008)
43. Forder, J., Allan, S.: Competition in the English nursing homes market. PSSRU Discussion Paper 2820, University of Kent (2011)
44. Grabowski, D.: A longitudinal study of medicaid payment, private-pay price and nursing home quality. Int. J. Health Care Financ. Econ. **4**(1), 5–26 (2004)
45. Starkey, K.B., Weech-Maldonado, R., Mor, V.: Market competition and quality of care in the nursing home industry. J. Health Care Financ. **32**(1), 67–81 (2005)
46. Gutacker, N., Bojke, C., Daidone, S., Devlin, N.J., Parkin, D., Street, A.: Truly inefficiency or providing better quality of care? Analysing the relationship between risk-adjusted hospital costs and patients' health outcomes. Health Econ. **22**, 931–947 (2013)
47. Mukamel, D.B., Spector, W.D.: Nursing home costs and risk-adjusted outcome measures of quality. Med. Care **38**(1), 78–89 (2000)
48. Laine, J., Linna, M., Häkkinen, U., Noro, A.: Measuring the productive efficiency and clinical quality of institutional long-term care for the elderly. Health Econ. **14**, 245–256 (2005a)
49. Laine, J., Linna, M., Noro, A., Häkkinen, U.: The cost efficiency and clinical quality of institutional long-term care for the elderly. Health Care Manag. Sci. **8**, 149–156 (2005b)
50. Weech-Maldonado, R., Shea, D., Mor, V.: The relationship between quality of care and costs in nursing homes. Am. J. Med. Qual. **21**(1), 40–48 (2006)
51. Mor, V., Morris, J., Lipsitz, L., Fogel, B.: Benchmarking quality in nursing homes: the Q-Metrics System. Can. J. Qual. Health Care **14**, 12–17 (1998)
52. Institute of Medicine.: Performance Measurement: Accelerating Improvement. The National Academies Press, Washington, DC (2006)
53. Abdi, H., Williams, L.J.: Principal component analysis. WIREs Comp. Stat. **2**, 433–459 (2010)
54. Jolliffe, I.: Principal component analysis. Encyclopedia of Statistics in Behavioral Science. Wiley, New York (2005)
55. Evans, R.: "Behavioural" cost functions for hospitals. Can. J. Econ. **4**, 198–215 (1971)
56. Kendall, M.G.: Rank Correlation Methods. Hafner Publishing Co, New York (1955)
57. Stock, J.H., Watson, M.W.: Heteroskedasticity-robust standard errors for fixed effects panel data regression. NBER Technical Working Paper 323 (2006)
58. Kezdi, G.: Robust standard errors estimation in fixed-effects panel models. Hung. Stat. Rev. Spec. **9**, 96–116 (2004)
59. Rogers, W.H.: Regression standard errors in clustered samples. Stata Tech. Bull. **13**, 19–23 (1994)
60. Clark, T.S., Linzer, D.A.: Should I use fixed or random effects? Polit. Sci. Res. Methods **3**(02), 399–408 (2015)
61. Cameron, A.C., Trivedi, P.K.: Microeconometrics Using Stata, Revised edn. Stata Press, College Station (2010)
62. Cameron, A.C., Trivedi, P.K.: Microeconometrics. Methods and Applications. Cambridge University Press, New York (2005)
63. Shea, J.: Instrument relevance in multivariate linear models: a simple measure. Rev. Econ. Stat. **79**(2), 348–352 (1997)
64. Hayashi, F.: Econometrics. Princeton University Press, Princeton (2000)
65. Bowblis, J.R., McHone, H.: An instrumental variables approach to post-acute care nursing home quality: Is there a dime's worth of evidence that continuing care retirement communities provide higher quality? J. Health Econ. **32**(5), 980–996 (2013)
66. Hirth, R.A., Grabowski, D., Feng, Z., Rahman, M., Mor, V.: Effect of nursing home ownership on hospitalization of long-stay residents: an instrumental variables approach. Int. J. Health Care Financ. Econ. **14**(1), 1–18 (2014)

67. Angrist, J.D., Krueger, A.B.: Instrumental variables and the search for identification: from supply and demand to natural experiments. J. Econ. Perspect. **15**(4), 69–85 (2001)
68. Drukker, D.M.: Testing for serial correlation in linear panel-data models. Stata J. **3**, 168–177 (2003)
69. Wooldridge, J.M.: Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge (2002)
70. Hahn, J., Ham, J.C., Roger Moon, H.: The Hausman test and weak instruments. J. Econom. **160**(2), 289–299 (2011)