

HEALTH VALUE ADDED OF HEALTHCARE ENTITIES*

Wiktor Ejsmont

University of Wrocław, Wrocław, Poland
e-mail: wiktor.ejmont@math.uni.wroc.pl
ORCID: 0000-0002-4516-9374

Janusz Łyko

Wrocław University of Economics and Business, Wrocław, Poland
e-mail: janusz.lyko@ue.wroc.pl
ORCID:0000-0002-0332-1921

© 2020 Wiktor Ejsmont, Janusz Łyko

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

Quote as: Ejsmont, W., and Łyko, J. (2020). Health value added of healthcare entities. *Econometrics. Ekonometria. Advances in Applied Data Analysis*, 24(3).

DOI: 10.15611/eada.2020.3.04

JEL Classification: A13, A14, I12, I13, I15, I18

Abstract: The paper deals with an evaluation of the quality of services provided by healthcare organizations. First, an index representing a patient's health condition is described, then its changes before and after being treated by a given entity are employed as a criterion to assess the operations of this entity. The index of a patient's health condition is based on the theory of survival analysis, while a model of random effects is used to determine the quality of services based on health value added.

Keywords: effectiveness, health value added, analysis of panel data, measurement of health.

1. Introduction

Quantitative research concerning health, healthcare and general medical issues have become more frequent recently. The current literature provides more and more evidence of such studies. A classic example of a book that presents all seminal research trends in this area is that by Shoukri and Cihon (1998). The authors present comprehensively statistical models applied by researchers of health in a broad understanding of this term.

One part of healthcare research deals with evaluating performance of healthcare entities, exemplified by studies using mostly logistic regression models. Logistic

* This work was supported by the Narodowe Centrum Nauki grant No 2018/29/B/HS4/01420.

regression is used in statistics when a variable under study takes two values. In the case of healthcare entities, the explained variable is equal to 1 or 0 depending on whether the patient was restored to health or not. Logistic regression allows to calculate the probability of the patient's recovery. This type of research is provided by Tay (2002) and Normand and Shahian (2007). Generally, the most popular statistical models employed to assess the performance quality of healthcare entities are discrete models, e.g. Geweke, Gowrisankaran and Town (2003), mostly due to the type of discrete data available for research on hospitals.

The paper aims to put forward a method of evaluating the performance quality of healthcare entities, i.e. a model that will determine the Health Value Added (HVA) based on the survival analysis theory and the random effects model. Health value added has been researched and reported by, e.g. Friedman, Kokia and Shemer (2003) who characterize it as a quantity representing a complex model of healthcare management that aims to maximize the quality of medical services. The authors claim that when measuring hospital performance one should consider not only the final outcomes, but also the process that contributed to those effects. It is much easier to treat a patient who is hospitalized at an early stage of illness than when the illness is more advanced. Therefore, it is necessary to find out how much has the hospital contributed to the patient's recovery, and to evaluate the hospital's performance against this background.

The idea to employ the random effects model is based on the concept of computing the patient's chance to survive at least one more year. Such a quantity can be computed before and after healthcare treatment. The difference between the two values represents the contribution of the healthcare entity to the improvement of the patient's health condition and may be used as a tool to evaluate the performance quality of the healthcare organization. When defining a patient's health condition, one can employ a survival analysis which is a well-known mathematical and statistical theory allowing to determine the duration of a patient's life.

2. Survival analysis

Mathematical survival theory aims at estimating the distribution of the survival function based on the characteristics of the population under study. The differences between the investigated individuals result from specific factors such as age, sex, smoking habits, etc. Certainly, other variables should also include some medical tests. Their inclusion should be preceded by consulting specialists, i.e. physicians, and also by testing for the statistical significance of each indicator. These factors, denoted by z_1^i, \dots, z_k^i for each individual $i \in \{1, 2, \dots, n\}$, represent the explanatory variables in a given model.

One may put forward several dozen different measurable medical factors that should be considered when assessing a medical condition. There are many more such characteristics and it is not possible to include all of them whilst estimating a survival

model. One example of the relevant factors is information about the patient’s previous accidents before being admitted to the hospital, that could significantly shorten his or her lifetime, irrespective of the actual medical conditions.

Assuming that T is nonnegative random variable, i.e. $P(T \geq 0) = 1$, which represents the duration of life, F and f denote the cumulative distribution function and the probability density function, respectively, of random variable T , one may define the survival function as:

$$S(t) = P(T \geq t) = 1 - F(t^-), \tag{1}$$

where t^- denotes the left-handed limit of the cumulative distribution function F at point t . $S(t)$ should be interpreted as a function indicating that an individual’s lifespan will be at least t years.

The survival function is related to the function

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{P(t \leq T < t + h | T \geq t)}{h}, \tag{2}$$

because density function f is continuous. This is the second important function in survival analysis – the hazard (intensity) function. One can demonstrate that

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \ln S(t) \tag{3}$$

and as a consequence

$$S(t) = \exp\left(-\int_0^t \lambda(x) dx\right), \tag{4}$$

obtain $S(0) = 1$ by assumption.

Cox (1972) introduced the Cox proportional hazard model in the form of

$$\lambda_i(t) = h(t) \exp\left(\sum_{l=1}^k \beta_l z_l^i\right), \tag{5}$$

where $h(t)$ is the non-negative baseline function describing the risk at time t if all factors are zero, i.e., $z_l^i = 0$. The survival function in the Cox model takes the form

$$S_i(t) = \exp\left(-\exp\left(\sum_{l=1}^k \beta_l z_l^i\right) \int_0^t h(x) dx\right) \tag{6}$$

and denoting $S_0(t) = \exp\left(-\int_0^t h(x) dx\right)$ one obtains

$$S_i(t) = S_0(t) \exp\left(\sum_{l=1}^k \beta_l z_l^i\right). \tag{7}$$

The form of function $h(t)$ does not matter when comparing the life durations of two individuals, because for any $i_1, i_2 \in \{1, 2, \dots, n\}$

$$\lambda_{i_1}(t) / \lambda_{i_2}(t) = \exp\left(\sum_{l=1}^k \beta_l (z_l^{i_1} - z_l^{i_2})\right) \quad (8)$$

and hence the model is sensitive to the variations of specific attributes z_1^i, \dots, z_k^i irrespective of $h(t)$. Assuming $h(t) = e^\alpha$ one can find the maximum likelihood estimates of $\hat{\alpha}, \hat{\beta}_1, \dots, \hat{\beta}_k$ using historical data, e.g. Hosmer and Lemeshow (1999), Magiera (2007).

One of the variables z_1^i, \dots, z_k^i represents the age of i -th individual expressed in months. Assuming z_1^i one can define the quantity

$$\pi^i = P(T \geq z_1^i + 12 | T \geq z_1^i) \quad (9)$$

representing the chances of individual i to survive at least one year given that i had already survived z_1^i months. Applying standard formulas for conditional probability, one obtains the following:

$$\begin{aligned} \pi^i &= P(T \geq z_1^i + 12 | T \geq z_1^i) \\ &= \exp\left(- (z_1^i + 12) \exp\left(\alpha + \sum_{j=1}^k \beta_j z_j^i\right)\right) / \exp\left(- z_1^i \exp\left(\alpha + \sum_{j=1}^k \beta_j z_j^i\right)\right) \\ &= \exp\left(-12 \exp\left(\alpha + \sum_{j=1}^k \beta_j z_j^i\right)\right). \end{aligned} \quad (10)$$

This paper applies the term health condition as a quantity expressed by formula (10) computed at two different times, i.e. when the individual enters and leaves hospital. The quantity can represent the health condition of i -th individual given health factors z_1^i, \dots, z_k^i before and after the medical treatment, e.g. at the time of entering hospital and at the time of leaving hospital. Hence by applying the random effects model, one can measure the contribution of the entity to increase the probability of the individual to survive at least one year, and consequently to determine the entity's health value added that will allow to assess the quality of the entity's services.

3. Random effects model

Before giving a more detailed description of the results the author introduced some notation and formulas. The model's denotations:

- x_{ij} – health condition of i -th patient at time of entering j -th hospital,
- y_{ij} – health condition of i -th patient at time of leaving j -th hospital,
- n_j – the number of patients at j -th hospital,
- m – the number of hospitals $j \in \{1, \dots, m\}$,
- n – the number of all the patients, i.e. $n = n_1 + \dots + n_m$,

\bar{x} – arithmetic average health condition of all the patients entering hospital,
 \bar{y} – arithmetic average health condition of all the patients leaving hospital,
 \bar{x}_j, \bar{y}_j – arithmetic average score for the health condition when entering and leaving, respectively, j -th hospital.

The proposed model can be applied for the assessment of the service quality of m homogenous entities of healthcare system, e.g. hospitals, hospital departments, etc. The homogeneity of services provided by healthcare entities is required for the sake of comparability. The author assumes that n_j out of n patients are served by j -th entity and their health condition before the treatment is defined by quantity $x_{ij} = \pi_{ij}$ (input health condition). After the treatment, a new value $y_{ij} = \pi_{ij}$ is computed based on new attributes z_1^i, \dots, z_k^i that now defines the patient's health condition.

As a result, one obtains unbalanced panel data since the number n_j of observations for individual hospitals can vary. When modelling an inhomogeneous population, one has to introduce inhomogeneity into the model. As regards the data under study, there may be various relations between the output and input variables for respective entities. The econometric literature refers to this model as the unbalanced one-way error component model with random effects, e.g. Baltagi (2005). The random effects model is also known as a variance components model (VC), cf. Maddala (2001). Wansbeek and Kapteyn (1982a, 1982b) first introduced this model (unbalanced panel data). The model takes the form:

$$y_{ij} = a + bx_{ij} + \xi_j + e_{ij}, \quad (11)$$

where e_{ij} is a random variable following a normal distribution $N(0, \sigma^2)$, whereas ξ_j follows $N(0, \sigma^2)$. In addition, it is assumed that random components from different entities and different patients are uncorrelated and that individual random term ξ_j is uncorrelated with random term e_{is} , i.e. $E(\xi_j, e_{is}) = 0$ for $j \neq s$ and $j \neq s$.

It follows from the form of the model that ξ_j represents a deviation of the average score for j th entity from the average score of the entire population. This average score in Figure 1 is shown as a dotted line, while a solid line illustrates the average score of the entire population. If ξ_j is positive then one may argue that j th entity improved its quality with respect to the average score of the entire population, while a negative ξ_j indicates the deterioration of its score compared to the average score of the population. Therefore, the value of parameter ξ_j is called value added or the operational effectiveness of the entity under evaluation. The value e_{ij} on the other hand represents the deviation of the individual patient's score from the average score of j -th entity where he/she was treated.

The above model is estimated by means of the maximum likelihood method (Aitken and Longford, 1986). The formulas for the estimates are given by Baltagi (2005) and Ejsmont (2009), where also a complete algorithm of estimating variance components σ^2 and σ_{ξ}^2 is provided. The obtained random effects are tested for significance by means of the Breusch-Pagan test (e.g. Baltagi, 2005).

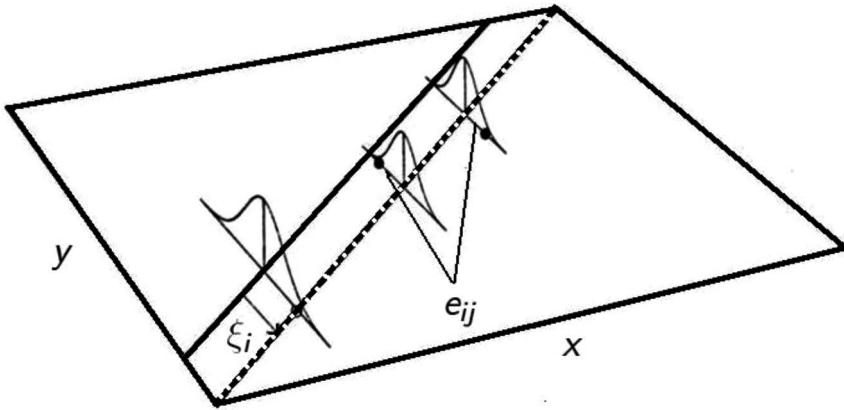


Fig. 1. The concept of measurement in a random effects model

Source: own elaboration based on (Skrondal and Rabe-Hesketh, 2008, p. 96).

In order to estimate the value of ξ_j , one can use the mean squared error theorem (Jakubowski and Sztencel, 2004, p. 135). Since both terms σ^2 and σ_I^2 are available before the model estimation, thus one can use them as the *a priori* information. Next, one determines the conditional distribution of random variable ξ_j given \bar{y}_j . The mean of *j*-th hospital has the form:

$$\bar{y}_j = a + b\bar{x}_j + \xi_j + \bar{e}_j \tag{12}$$

and under appropriate assumptions it is distributed as $N(a + b\bar{x}_j, \sigma_I^2 + \sigma^2 / n_j)$. Since ξ_j follows $N(0, \sigma_I^2)$, thus the conditional distribution $f(\xi_j / \bar{y}_j)$ is normal and has the form:

$$N(\rho n_j^*(\bar{y}_j - a - b\bar{x}_j), n_j^*(1 - \rho)\sigma_I^2 / n_j), \tag{13}$$

where $\rho = cor(y_{ij}, y_{pj}) = \sigma_I^2 / (\sigma_I^2 + \sigma^2)$, $n_j^* = w_j / (1 - \rho)$ and $w_j = n_j\sigma^2 / (\sigma^2 + n_j\sigma_I^2)$. As a result, a comparison of the operational quality of healthcare entities will be based on comparing the mean values from the conditional distribution, i.e.

$$e_j = \hat{\rho}n_j^*(\bar{y}_j - \hat{a} - \hat{b}\bar{x}_j), \tag{14}$$

where $\hat{\rho}, \hat{a}, \hat{b}$ respectively are estimates of: correlation ρ , intercept a and slope b .

4. Conclusion

The presented model enables the evaluation of the operational quality of healthcare entities. The evaluation is relative and shows the position of a given entity against

the background of the average performance of other similar organizations. Consequently, the ranking of entities providing similar healthcare services can be obtained. It is worth emphasizing that the model includes both the direct effect of the treatment, i.e. whether the patient recovered or not, and also the entity's contribution to the recovery of the patient. Full recovery from a serious health hazard condition is evaluated differently than the provision of standard treatment when the patient's general state of health is good. By monitoring the patient's current health condition with certain medical characteristics that determine one indicator, i.e. the probability of the patient surviving one year, it is possible to determine by how far the provision of the medical service at a given entity contributed to the change of this index. By averaging those quantities for each entity under evaluation, one can determine health value added that is generated at this entity. The article is theoretical but it can be used to assess health value added in terms of data.

References

- Aitken, M., and Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society*, 149(1), 1-43.
- Baltagi, B. H. (2005). *Econometric analysis of panel data*. John Wiley & Sons Ltd.
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2), 187-220.
- Ejsmont, W. (2009). Efektywność nauczania we wrocławskich liceach. *Didactics of Mathematics*, 5-6(9-10), 111-128.
- Friedman, N. L., Kokia, E., and Shemer, J. (2003). Health Value Added (HVA): Linking strategy, performance, and measurement in healthcare organizations. *Israel Medical Association Journal*, 5(2), 3-8.
- Geweke, J., Gowrisankaran, G., and Town, R. (2003). Bayesian inference for hospital quality in a selection model. *Econometrica*, 71(4), 1215-1238.
- Hosmer, D. W., Jr., and Lemeshow, S. (1999). *Applied survival analysis. Regression modeling of time to event data*. John Wiley & Sons.
- Jakubowski, J., and Sztencel, R. (2004). *Wstęp do teorii prawdopodobieństwa*. Warszawa: SCRIPT.
- Maddala, G. S. (2001) *Introduction to econometrics*. John Wiley & Sons.
- Magiera, R. (2007). *Modele i metody statystyki matematycznej*. Wrocław: Oficyna Wydawnicza GiS.
- Normand, S. L., and Shahian, D. (2007). Statistical and clinical aspects of hospital outcomes profiling. *Statistical Science*, 22(2), 206-226.
- Shoukri, M. M., and Cihon, C. (1998). *Statistical methods for health sciences*. CRC press.
- Skrondal, A., and Rabe-Hesketh, S. (2008). *Multilevel and longitudinal modeling using stata*. College Station, Texas: Stata Press Publication – StataCorp LP.
- Tay, A. (2002). *What can patient outcomes tell us about hospital quality?* (Working Paper, Columbia University, USA, Dept. of Economics).
- Wansbeek, T. J., and Kapteyn, A. (1982a). A class of decompositions of the variance-covariance matrix of a generalized error components model. *Econometrica*, 50(3), 713-724.
- Wansbeek, T. J., and Kapteyn, A. (1982b). A simple way to obtain the spectral decomposition of variance components models for balanced data. *Communications in Statistics*, (11), 2105-2112.

ZDROWOTNA WARTOŚĆ DODANA JEDNOSTEK SŁUŻBY ZDROWIA

Streszczenie: W artykule został podjęty temat oceny jakości usług jednostek służby zdrowia. W pierwszej kolejności opisano wskaźnik służący do określenia zdrowotnej kondycji pacjenta, a następnie jego zmiany w okresie przed skorzystaniem z usługi i po skorzystaniu z niej. Użyto go jako kryterium oceny pracy danej jednostki. Wskaźnik oceniający stan zdrowia pacjenta skonstruowano, wykorzystując teorię analizy przeżycia, a do określenia jakości usług na podstawie zdrowotnej wartości dodanej posłużono się modelem efektów losowych.

Słowa kluczowe: efektywność, zdrowotna wartość dodana, analiza danych panelowych, pomiar zdrowia.