Efficiency analysis of public health spending in Brazilian capitals using network Data Envelopment Analysis

Maria Stella de CASTRO LOBO

Rio de Janeiro Federal University UFRJ, Brazil

Edson CORREIA ARAUJO

World Bank, USA

Abstract:

Aim: In 1988, Brazil implemented profound changes in the organization and financing of its public health system, with the creation of the Unified Health System (Sistema Unico de Saúde – SUS), establishing universal health coverage. The gradual expansion of the health system and entitlements to services has been accompanied by the debate about the appropriate level of government spending and health system efficiency.

Design / Research methods: The study uses Variable Returns to Scale output-oriented, Dynamic Network Slacks-Based-Measure Data Envelopment Analysis (DEA) model, period 2008-2013, to depict the relationships that take place between diverse levels of care (primary health care/PHC and secondary-tertiary health care/STC). Decision Making Units are Brazilian state capitals, which implement key health policies and assist patients from smaller surrounding municipalities, especially for STC. Inputs are PHC and STC budgets; outputs are their respective services provided and avoidable deaths. The link variable is PHC medical consultation, entrance door to the system and gatekeeper for more complex levels of care. Dynamic model evaluates efficiency across time.

Conclusions / findings: Overall performance was 0.86; for PHC, 0.90; for STC, 0.85 (SD=0.15). 8 out of 27 capitals were fully efficient. Capitals increased average scores in both levels of care, but only STC had a positive technological change (frontier shift >1). Link variable behavior denotes a bottleneck between levels of care. Projections onto the frontier enable establish own management diagnosis and goals for financing and development.

Originality / value of the article: Network models mimic hierarchically organized health systems. The appliance of results aids health policy.

Keywords: network DEA, efficiency measurement, health and economic development, public health
JEL: H21, H51, O10, O54

1. Introduction

Brazil covers 8.5 million km², or 47% of South America, with an estimated population of 204,482,509 in 2015. As a federative republic, its political organizational system has three levels of autonomous government—federal government, 26 states and a federal district (federal capital), and 5,570 municipalities. The country is divided into five geographical regions (north, northeast, center-west, southeast, and south) with differing demographic, economic, social, cultural, and health conditions, and widespread internal inequalities. The north region, which contains most of the Amazon rainforest, has the country's lowest population density (3.9 people per km^2) and is the second poorest region, after the northeast region.

Since 1970, Brazil underwent a demographic transition: urban population increased (from 55.9% to 84.0%); fertility rate and infant mortality rate decreased (respectively, from 5.8 to 1.8 and from 114.0 per 1,000 livebirths to 15.3 per 1,000 livebirths). The population older than 60 years doubled (11.0% in 2012) and life expectancy at birth increased to 74.8 years, resulting in epidemiological transition. Nowadays, diseases of the circulatory system are the leading cause of death, followed by cancer and external causes (mainly traffic accidents and homicides). Chronic diseases are the biggest contributor to the burden of disease (24.3% of adult population have hypertensive disorders; 11.7% are diabetic), although communicable diseases still affect a substantial proportion of the population, especially in the poorest areas in the country (Ministry of Health: DATASUS 2017).

Over the last decades, Brazil implemented profound changes in the organization and financing of its public health system. The creation of the Unified Health System (*Sistema Unico de Saúde* – SUS) in 1988 Constitution, establishing universal health coverage, has been associated with the expansion of the health service delivery system with remarkable improvements in access, financial protection and health outcomes (Gragnolati et al. 2013: 16).

The SUS main principles are universal coverage and equity of care; its main guidelines are decentralization, community participation and comprehensive care,

which means continuous provision of health services at all levels of complexity (from preventive actions to curative high technology procedures) (Paim et al. 2011: 1787).

Primary health care (PHC) is the main entrance door to the system. PHC aims to provide universal access and comprehensive health care, coordinate and expand coverage to more complex levels of care (specialist care and hospital care, that is, STC or secondary – tertiary levels of care), and implement intersectoral actions for health promotion and disease prevention.

Since its creation, the SUS network expanded considerably, particularly outpatient services driven by the expansion of the family health strategy (*FHS*), to guarantee comprehensive primary care. Between 1998 and 2015, FHS teams increased from 4,000 to over 50,000, covering 56.2% of the population (Malta et al. 2016: 328). From 2008 to 2015, the percentage of STC admissions due to conditions sensible to primary care fell from 35.8 % to 30.6%.

By decentralization, municipalities took on a leading role in the delivery of health services, while states and federal government maintained responsibility for some referral services. Tripartite councils – municipal, state and federal – sign commitments to health goals and guarantee shared financing responsibility, but municipalities still define priorities, set goals with local health services and allocate the final budget according to management contracts.

Budget transfer to municipalities considers six financing blocks or sub-functions: 1) Primary Care (PHC); 2) Medium/ High Complexity Outpatient Hospital Care (Secondary – Tertiary Care – STC); 3) Health Surveillance, 4) Pharmaceutical Care, 5) SUS Management, 6) Investments on Healthcare Services Network. By 2010, primary care and outpatient /hospital care consumed 14.3% and 52.0%, respectively, from total federal public health spending. (Paim et al. 2011: 1782).

The gradual expansion of the health system and entitlements to services has been accompanied by the debate about the appropriate level of government spending and health system efficiency. Total health expenditures in Brazil is comparable to the average of the Organization for Economic Cooperation and Development (OECD) countries; Brazil spends 8.3 % (2015) of its gross domestic product (GDP)

on health while the average for OECD countries is 9.0 % (2015). However, health spending in Brazil is dominated by private insurance, which accounts for more than half of the health expenditure, and covers approximately a quarter of the population who pays for it. This results in large inequalities in the per capita spending between private and public sector (USD 2,678 at PPP versus USD 1,028 at PPP). Additionally, health indicators in Brazil are still below the average among OECD countries and, in many cases, worse than its regional and economic peers. In terms of health services delivery indicators, such as consultations per physician or bed occupancy rates, Brazil lags behind comparable countries (OECD and regional and structural peers) (OECD 2015: 14). Despite its considerable achievements, the SUS still faces vital challenges across all levels of care. There remain significant coverage gaps in PHC, barriers to access specialist and high-complexity care (STC), weak quality and coordination of care, and weaknesses in the referral and counter-referral systems (Gragnolati et al. 2013: 66).

Altogether, despite the efforts to establish a public universal health system, Brazil struggles to achieve a good balance between an appropriate level of (public) spending and to obtain better value for the resources invested in its health system. The objective of this study is to measure efficiency in public health spending in a model that considers efficiency´s operational (delivery of health services) and quality (health indicators) dimensions, and the interconnectedness between levels of care, which systemically influence the final results.

2. Methods: conceptual framework for current analysis

Health care systems (HCS) have great impact on society, demand complex decision making and operational research (OR) can provide useful tools to help managers (Alexander 2008: 97). HCS structure is multidimensional, hierarchically organized, demands communication skills to understand it in depth and presents gaps of data, sometimes being difficult to formulate and define boundaries to model. In operational terms, to analyze the health system as a whole, it is important to consider not only the interconnectivity between components of the system (levels of

care), but also the diverse perspectives of stakeholders (quantitative – volume of health services delivery; qualitative – health indicators) and the influence of environment (socioeconomic influences) (Emrouznejad, de Witte 2010: 1574-1575).

Moreover, with always expanding incorporation of new technologies and divergent market interests guiding regulation, healthcare services consume increasing proportions of GDP, not necessarily associated with quality of care. For these reasons, OR efficiency and performance studies have been frequently chosen to handle the problem and monitor public expending (Lobo, Lins 2010: 380).

Data Envelopment Analysis (DEA) is a non-parametric technique which allows to construct a data driven production frontier. The estimated best practice efficiency frontier represents the maximum level of outputs given the inputs and the technology available. By comparing similar decision making units (DMU), the technique identifies the DMUs with higher level of production for fixed amounts of inputs (output-oriented model) and/or the ones that use fewer resources to generate a fixed amount of products (input-oriented model). Generally, DEA production frontiers can be either constant returns scale (CRS) or variable returns to scale (VRS). To be fully efficient, a DMU should be located in Pareto-efficient portions of the frontier, that is, a place where it is not possible to reduce any input, or increase any output, without having to also increase another input or reduce another output, simultaneously (Cooper et al. 2007: 1-39, 45-46).

The traditional DEA model is often considered a black-box model, as it specifies what enters and gets out of the transformation process (exogenous inputs and outputs) without explicitly modelling what happens inside it. To overcome these challenges, an extension of the DEA black box model is the network DEA model (Färe, Grosskopf 2000: 35).

The network model attempts to analyze inside of a production process and to explicitly model the relationships among variables that take place inside the DMU. For example, when dealing with more than one dimension, division or sub-process inside the black box, there are connections among them and there may exist link variables present in more than one dimension, which function as inputs for one dimension and as outputs for another. Network model allows to calculate the total efficiency score and specific efficiency scores, to define benchmarks for each sub-

process, and recommend for each variable to project onto the best practice frontier (Tone, Tsutsui 2009: 244-246).

Selecting different inputs and outputs in DEA models can heavily influence the results; the same happens when arranging variables and relationships inside the black box. In other words, the architecture of the relationships inside the black box influence the efficiency scores and possibilities for projections onto the best practice frontier. This calls for the importance of the managerial analysis and close interaction with stakeholders when designing accurate models (Emrouznejad, de Witte 2010: 1579).

A DEA dynamic network model allows to incorporate changes in the efficiency score across years, by incorporating the existence of time intermediate carry-overs that connect two consecutive terms. Longitudinal analysis uses a modified Malmquist Index, that evaluates individual DMU score changes across time (catch-up component) and the technological or frontier dislocation of all units on the analyzed period (frontier shift component). Finally, to cope only with Pareto-efficient projections, a network slacks-based measure (NSBM) approach is proposed. Slack-based measure is a non-radial method and is suitable for measuring efficiencies when inputs and outputs may change non-proportionally, turning recommendations more reliable (Tone 2001).

2.1 Dynamic Network SBM DEA Model: mathematical modeling

This paper utilizes network SBM (non-radial) DEA model, as proposed by Tone and Tsutsui (2009), in which the definition of efficiency scores of each observed DMU depends on the selected orientation, input, output or non-oriented. Thus, we work with output-oriented SBM network model under the assumption of VRS and free link case. It considers a system of two sub-processes or divisions, linked by one intermediate product from process one to process two.

To formalize, we deal with n DMUs (j=1, …, n) consisting of K divisions (k=1, …, K). Let m_k and r_k be the numbers of inputs and outputs to Division k, respectively. We denote the link leading from Division k to Division h by (k,h). ρ^*_o denotes the overall efficiency of DMU$_o$; s_r^{k-} and s_r^{k+} are input and output slacks of process K.

The production possibility set $\{(x^K, y^k, z^{(k,h)})\}$ is defined by:

$$x^k \geq \sum_{j-1}^{n} x_j^k \lambda_j^k \quad (k = 1, \ldots, K$$

$$y^k$$
$$\leq \sum_{j-1}^{n} y_j^k \lambda_j^k \quad (k$$
$$= 1, \ldots, K) \tag{1}$$

$$z^{(k,h)} = \sum_{j-1}^{n} z_j^{(k,h)} \lambda_j^k \quad (\forall(k,h)) \qquad \text{(as outputs from } k\text{)}$$

$$z^{(k,h)} = \sum_{j-1}^{n} z_j^{(k,h)} \lambda_j^h \quad (\forall(k,h)) \qquad \text{(as inputs to } h\text{)}$$

$$\sum_{j-1}^{n} \lambda_j^k = 1 \ (\forall k), \qquad \lambda_j^k \geq 0 \ (\forall j, k)$$

where λ_j^k is the intensity vector corresponding to Division k (k=1, ..., K).

DMUo (o= 1, …, n) can be represented by:

$$x_o^k = X^k \lambda^k + s^{k-} \quad (k = 1, \ldots, K),$$
$$y_o^k = Y^k \lambda^k + s^{k+} \quad (k = 1, \ldots, K), \tag{2}$$
$$e\lambda^k = 1 \ (k - 1, \ldots, K),$$
$$\lambda^k \geq 0, s^{k-} \geq 0, s^{k+} \geq 0, \qquad (\forall k),$$

And

$$X^k = \left(x_1^k, \ldots, x_n^k\right) \in \mathbb{R}^{m_k \times n} \tag{3}$$
$$Y^k = \left(y_1^k, \ldots, y_n^k\right) \in \mathbb{R}^{r_k \times n}$$

As regard to the linking constraints, considering the ''free'' link value case, linking activities are discretionary and link flow may increase or decrease in the optimal solution:

$$Z^{(k,h)} \lambda^h = Z^{(k,h)} \lambda^k \ \left(\forall k(k,h)\right), \tag{4}$$

$$Z^{(k,h)} = \left(z_1^{(k,h)}, \ldots, z_n^{(k,h)}\right) \in \mathbb{R}^{t_{(k,h)} \times n} \tag{5}$$

We evaluate the output-oriented overall efficiency of DMUo (τ_o^*) by solving the following linear program:

$$\frac{1}{\tau_o^*} = \max_{\lambda^k s^{k+}} \quad \sum_{k=1}^{K} w^k \left[1 + \frac{1}{r_k} \left(\sum_{r=1}^{r_k} \frac{s_r^{k+}}{y_{ro}^{k+}} \right) \right] \tag{6}$$

Where $\sum_{k=1}^{K} w^k = 1; w^k \geq 0 \ (\forall K)$ and w^k is the relative weight of Division k which is determined corresponding to its importance (in this paper, $w_1 = w_2 = 0.5$).

In order to confine the score into the range [0, 1], we define the output-oriented divisional efficiency score by:

$$\tau_k = \frac{1}{1+\frac{1}{r_k}\left(\sum_{r-1}^{r_k}\frac{s_r^{k=1}}{y_{r_0}^k}\right)} \ (k = 1, \dots, K), \tag{7}$$

Where s^{k+} is the optimal output-slacks for (6). The output-oriented overall efficiency score is the weighted harmonic mean of the divisional scores:

$$\frac{1}{\tau_0^*} = \sum_{k-1}^{K}\frac{W_k}{\tau_k}. \tag{8}$$

In Dynamic slacks-based DEA model, intertemporal efficiency change considers carry-over activities (in this study, treated as fixed, non-discretionary variables). The objective function for output-oriented dynamic model is an extension of the output-oriented SBM network model and deals with shortfalls in output products and desirable (good) links because both have similar features, i.e. larger amount is favorable. As it produces a modified Malmquist Productivity Index (MPI), we bring the equation from the later, where y represents the output vector that can be produced using the input vector x. Two of the four distance functions, $D_t \ (x_t, \ y_t)$ and $D_{t+1} \ (x_{t+1}, \ y_{t+1})$, are technical efficiency measures in times t and t+1, respectively, and the remaining functions, $D_t \ (x_{t+1}, \ y_{t+1})$ and $D_{t+1} \ (x_t, \ y_t)$, indicate cross-period distance functions. $D_t \ (x_{t+1}, \ y_{t+1})$ shows the efficiency measure using the observation in time t+1 relative to the frontier technology in time t. $D_{t+1} \ (x_t, \ y_t)$ shows the efficiency measure using observation in t relative to the frontier technology in time t+1 (Färe at al. 1992, 1994).

$$\text{MPI}=\left[\frac{D_t(x_{t+1}, y_{t+1})}{D_t(x_t, y_t)} \times \frac{D_{t+1}(x_{t+1}, y_{t+1})}{D_{t+1}(x_t, y_t)}\right]^{1/2} \tag{9}$$

An important feature of the DEA based Malmquist index is that it can decompose the overall productivity measure into two mutually exclusive components (10), one measuring change in technical efficiency (catching-up effect, outside the brackets) and the other measuring change in technology (frontier shift or innovation, inside), with a cut-off point set at the unit for progression or regression across time.

$$\text{MPI} = \frac{D_{t+1}(x_{t+1}, y_{t+1})}{D_t(x_t, y_t)} \left[\frac{D_t(x_{t+1}, y_{t+1})}{D_{t+1}(x_{t+1}, y_{t+1})} \times \frac{D_t(x_t, y_t)}{D_{t+1}(x_t, y_t)} \right]^{1/2} \qquad (10)$$

2.2 Empirical model for Brazilian public health spending analysis: DMUs, variables and data sources

For public health spending analysis, DMUs are the Brazilian state capitals, the core and main municipalities of each state. They implement key health policies, operate service delivery and have high socio-political influence over surrounding municipalities of the same state, frequently receiving patients from the latter, especially for secondary-tertiary (STC) care.

As already mentioned, the analysis applies a VRS output-oriented, Dynamic Network SBM DEA (DNSBM) model, and considers period 2008-2013. The output-oriented model is chosen given the ultimate objective is to maximize outputs, that is, increase operational productivity and ameliorate health indicators, given the resources provided. The VRS model is justified since state capitals are different in scale. The two sub-processes or divisions are the PHC and STC levels of care, which are linked by one free intermediate link: PHC medical consults (output for PHC and input for STC).

In order to understand the architecture of the model, Figure 1 presents DNSBM model with two independent exogenous inputs: (i) total health expenditure on PHC; and (ii) total health expenditures on STC; the width of the latter arrow depicts larger quantity of money in Brazilian local currency (BRL). The model uses two health indicators as final exogenous outputs: (i) the number of avoidable deaths in the population of 0-4 years-old; and (ii) the number of avoidable deaths in the population aged between 5-74 years-old (again, the difference in width of the arrow

shows different volume of deaths, higher for adults and occurring inside hospitals). The avoidable or reducible causes of death are defined as those totally or partially preventable by health service effective actions, accessible in a determined place and time. The list is periodically revised worldwide by local committees of experts according to the International Code of Diseases (ICD). In Brazil, these causes were reviewed considering the available knowledge and technology for the practice of health care by a working-group organized by the Brazilian Ministry of Health, 2008. They are mostly associated with vaccination programs; prenatal, delivery, neonatal care; opportune diagnosis and treatment; health promotion and prevention for chronic diseases (Malta et al. 2010).

Considering that avoidable causes of death are undesirable final outputs, survival rates (complement of mortality) were used to maintain output-oriented maximizing model, as proposed by Afonso and Aubyn (2005: 238). Since mortality is the usual epidemiologic indicator, the data is reversed again and presented as mortality in results section in order to be consistent with the public health audience. That is why the output projection for mortality related outputs appears negative.

Figure 1. DEA SBM Dynamic Network Model for public health spending analysis
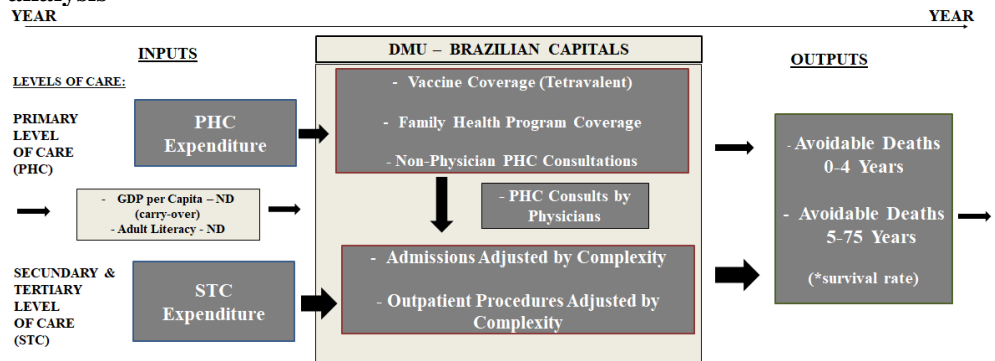
Source: Authors' own study

To understand how the public health care funds are allocated to purchase health services at different levels of care, it is necessary to explore the health care production process within each DMU. The production process within each DMU has two sub-processes or divisions: primary care (PHC) and secondary-tertiary services

(STC). PHC usually comprises appointments with physicians; primary dental care; primary procedures with other types of professionals, with medium or high degrees of skill; ambulatory and home assistance of FHP; vaccination; educational activities to community groups; pre-natal assistance; family planning activities; minor surgeries; activities of community workers; emergency help in basic unit of health. For STC, either medium or high complexity, there are outpatient and inpatient procedures (Varela et al. 2009: 115). Outpatient procedures comprise diagnostic and therapeutic activities, clinical or surgical, performed in ambulatory settings (in Brazilian time series, 15% of them consist of specialized consults). When admissions are necessary, inpatient hospital procedures occur at hospital level.

Inside the model, each dimension has its specific set of outputs. For PHC, they are: number of doses of tetra vaccine, number of persons registered in Family Health Program (FHP), number of non-medical primary care consultations. For STC, they are: number of admissions adjusted by complexity; number of outpatient procedures adjusted by complexity. Finally, the variable 'number of medical PHC consultations' links these two sub-processes, since PHC physicians play a role of gatekeepers, referring patients to STC level as needed. At STC level, the main point of contact is the specialized consult, unless an urgent admission is necessary. Inside the model, PHC consultations are outputs of PHC and inputs for secondary-tertiary care. Note the arrow-out means output, and arrow-in means input for a given level or sub-process.

The dynamic component to compare the frontier shift and observe the influence of specific policies on the DMUs performance across the years uses two fixed non-discretionary carry-over variables, namely: number of literate people over 15 years-old and GDP per capita. By definition, carry-over variables in dynamic network DEA are outputs in a given year and inputs for subsequent year. The sociodemographic variables control for socio-demographic diversities. They are both influenced by health status and health practices (output of the previous year), and also influence health status and health practices in the near future (input for the next year), with impact on performance scores.

Table 1. Variables and sources according to levels of care and behavior inside the model

Level		Variables – Characteristics (exogenous, inside DMU, carry-over)	Sources***
PHC	**Inputs**	Total public spending on PHC – EXOGENOUS	SIOPS
	Outputs	Avoidable causes of death mortality, 0-4 years-old * (outside hospital) – EXOGENOUS	SIM
		Avoidable causes of death mortality, 5-75 years-old *(outside hospital) – EXOGENOUS	SIM
		Number of administered doses of tetra or pentavalent vaccine – INSIDE DMU	SI- PNI
		FHP Coverage (number of persons) – INSIDE DMU	SIAB
		Number of primary care consultations (except physicians) – INSIDE DMU	SIA/SUS
		Number of primary care consultations (by physicians) – INSIDE DMU – LINK to STC	SIA/SUS
	Non-discretionary	GDP per capita – CARRY – OVER	IBGE
		Adult Literacy - CARRY – OVER	IBGE
STC	**Inputs**	Total public spending on STC - EXOGENOUS	SIOPS
	Outputs	Number of primary care consultations (by physicians) – INSIDE DMU - LINK to STC	SIA/SUS
		Number of Admissions Adjusted by Complexity ** – INSIDE DMU	SIH/SUS
		Number of Outpatient Procedures Adjusted by Complexity ** – INSIDE DMU	SIA/SUS
		Avoidable causes of death mortality, 0-4 years-old * (inside hospital) – EXOGENOUS	SIM
		Avoidable causes of death mortality, 5-75 years-old *(inside hospital) – EXOGENOUS	SIM
	Non-discretionary	GDP per capita – CARRY – OVER	IBGE
		Adult Literacy – CARRY – OVER	IBGE

Source: Authors' own study

* Undesirable final outputs treated as the complement (survival rates, as in Afonso, Aubyn 2005: 238). In Brazil, almost 85% of registered deaths occur inside hospitals.

**For hospital care, the adjustment factor was 6.0 for high complexity procedures (medium complexity admissions are 13.8 times more frequent while the system spends only 2.3 more compared to high complexity). For outpatient care, the adjustment factor was 1.1.

***Sources: SIOPS – Information System for Public Budgets for Health (Ministry of Health – MoH), SIAB – Information System for Primary Care – MoH, SI- PNI – Information System for the National Immunization Program – MoH, SIA/SUS and SIH/SUS, respectively, Information System for Ambulatory Care and Information System for Hospital Care – MoH, SIM – Information System for Mortality – MoH, IBGE – Geography and Statistics Brazilian Institute – Census Data.

Table 1 presents the list of variables and sources included in the model. Most sources or databases are administered by the Ministry of Health. As for Brazilian capitals, with robust administrative structure, data quality has not been a concern (that would be different if DMUs were the smaller or less equipped municipalities). Given accurate data, there was no need to create confidence intervals.
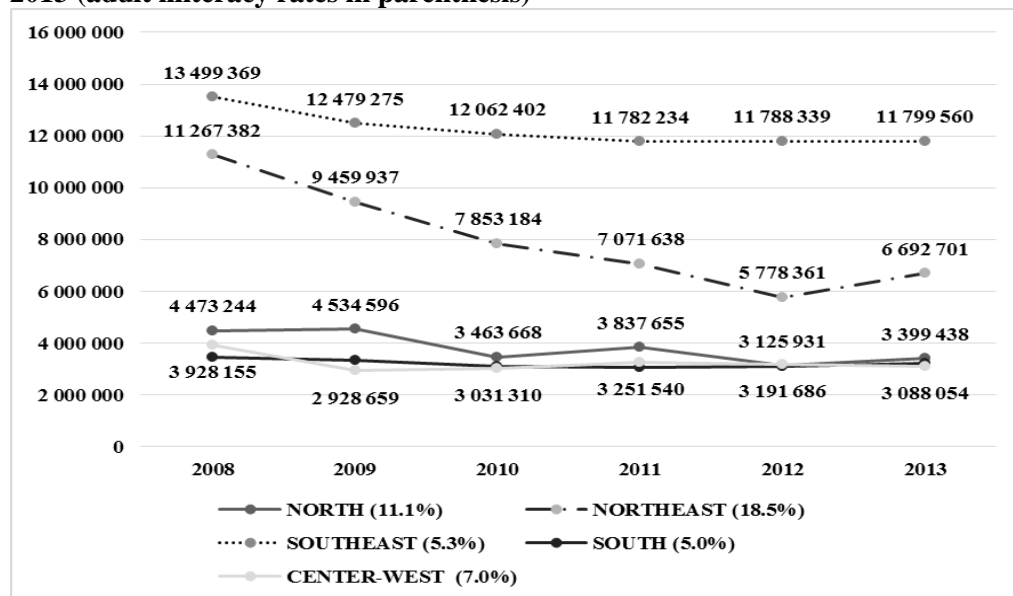
The variables were preferably treated as absolute numbers for two reasons (except for non-discretionary GDP per capita): first, the DEA literature argue that the use of ratios may damage the linear properties of the frontier, especially the convexity assumption (Emrouznejad, Amin 2009: 486). Second, in order to give recommendations based on projections onto the best practice frontier, absolute numbers seem more appropriate.

3. Results

From 2008 to 2013, there was an increase in both levels of expenditure, PHC spending increased by 66.2% and STC increased by 51.6%. At PHC level, FHP coverage, vaccination and non-medical consultations increased by 8.4%, 4.0% and 5.0%, respectively. Nevertheless, PHC medical consultations decreased by 23.0% in the period. At STC level, there was 32.1% and 12.9% increase of outpatient procedures and hospital admissions, respectively. Finally, there was a decrease in avoidable deaths for children below 5 years old (by 8.7%). For people 5-74 years-old, avoidable deaths increased by 4.4%. (Table 2)

Note that decreasing proportions of PHC medical consults were diverse among regions, which may reflect the difficulties in attracting physicians to work at this level. Figure 2 shows the differences in the number of PHC medical consultations across regions; the decline being sharper in the Northeast region, which complies a more challenging sociodemographic scenario, with the worse adult illiteracy rates (18.5% versus Brazilian average 9.4%).

Figure 2. Brazilian capitals´ average PHC medical consults by regions: 2008-2013 (adult illiteracy rates in parenthesis)

Source: Authors' own study

Table 2. Descriptive statistics of data, 2008-2013

YEAR	PHC Expenditure (R$ mi)	Avoidable Deaths 0-4 yo	Avoidable Deaths 5-74 yo	FHP (1,000)	Tetravalent Vaccine	Nonmedical Consults	Medical Consults	STC Expenditure (R$ mi)	Avoidable Deaths 0-4 yo	Avoidable Deaths 5-74 yo	Outpatient adjusted (1,000)	Admissions adjusted
		PRIMARY CARE					LINK		SECONDARY-TERTIARY			
2008												
Average	46.81	76	1,993	855	72,610	575,065	1,356,317	512.35	687	5,979	24,505	114,564
Max	281.32	300	10,798	5,789	473,723	3,243,461	8,324,869	3,179.42	2,699	32,393	174,742	564,640
Min	10.90	14	190	134	11,505	38,432	97,796	40.18	127	569	2,165	12,982
StDev	54.22	61	2,425	1,070	93,051	646,580	1,575,830	635.11	545	7,275	32,891	111,278
2013												
Average	77.78	70	2,081	926	75,527	603,719	1,043,997	776.90	627	6,243	32,375	129,361
Max	491.38	274	11,441	6,126	520,912	3,158,932	6,403,400	4,680.08	2,470	34,322	231,198	666,210
Min	11.62	13	253	99	13,639	46,544	168,311	72.28	115	758	2,343	26,427
StDev	99.93	56	2,473	1,202	102,984	732,955	1,249,480	920.98	504	7,420	44,230	129,924
Change (2013/2008)	66.2%	-8.7%	4.4%	8.4%	4.0%	5.0%	-23.0%	51.6%	-8.7%	4.4%	32.1%	12.9%

Source: Authors' own study

The average overall efficiency score for all Brazilian capitals was 0.86 (SD= 0.15). Figure 3 presents the overall, PHC and STC efficiency scores for each capital in the analyzed period. The continuous grey line represents the overall score and eight efficient capitals are fully efficient at both, PHC and STC levels of analysis. All capitals from the South – S region, three out of four capitals in Southeast region – SE, one out of five in North region - N and the capital of Brazil (federal district), in Center-West - CW, were efficient. None in the Northeast – NE was fully efficient.

For the remaining inefficient capitals, 11 had superior performance on PHC activities (mainly from N and NE); 8, on STC activities. In the case of Rio de Janeiro – SE, for instance, the capital was a hundred percent efficient at PHC level, but was inefficient at STC level (STC score equals 0.85). The opposite happened to Fortaleza – NE, 100% efficient at STC level, but inefficient at PHC (PHC score equals of 0.91). Twelve capitals were inefficient at both levels. The minimum observed efficiency score was 0.48 (Cuiabá – CW; 0.46 in PHC and 0.56 in STC).

Given the model is dynamic, there were variations across the years (Figure 4). In short, the panel nature of the data allowed the computation of productivity growth for all capitals over the period of 2008-2013. Overall performance increased from 0.86 (SD=0.16) to 0.89 (SD=0.15); PHC, from 0.90 (SD=0.17) to 0.94 (SD=0.14); STC level, from 0.84 (SD=0.18) to 0.88 (SD=0.19). Although there was approximation to the frontier, with ascending scores and catch-up average values above the unit, a positive rate for frontier shift or technological change occurred only for STC, by 3.0%. The STC technological change presented difference among regions according to sociodemographic gradient: South had the best performance (positive rate 6.1%), followed by Southeast (2.8%), North (2.2%). Northeast and Center-West had negative rates (-0.8% and -3.8%, respectively). Overall Malmquist stayed nearly unchanged in the period (=1.008). (Figure 5)
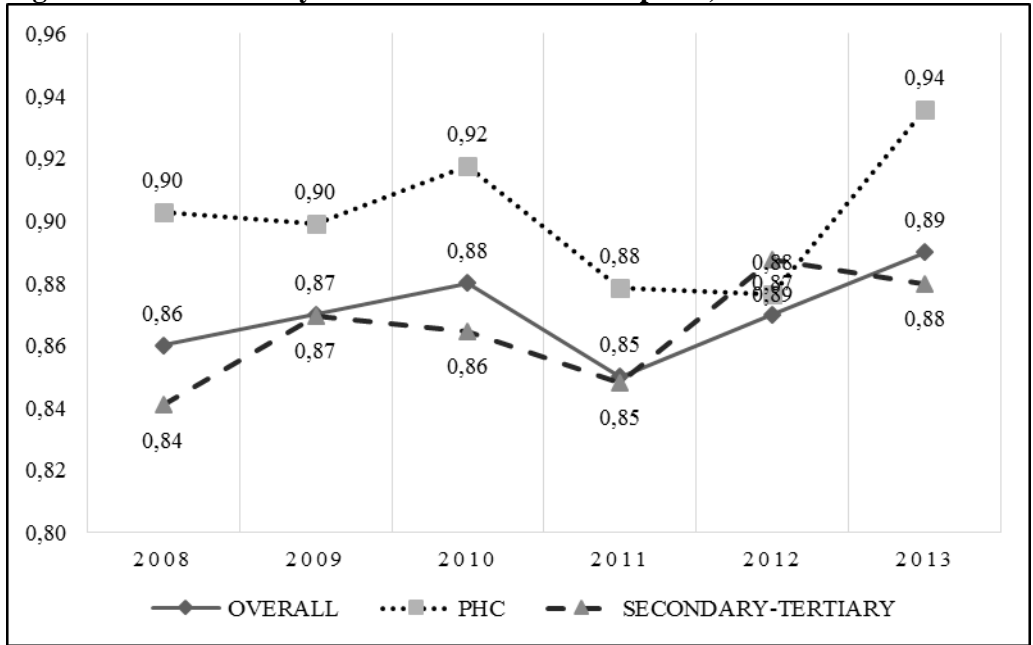
Figure 3. Efficiency scores - Brazilian state capitals, 2008-2013
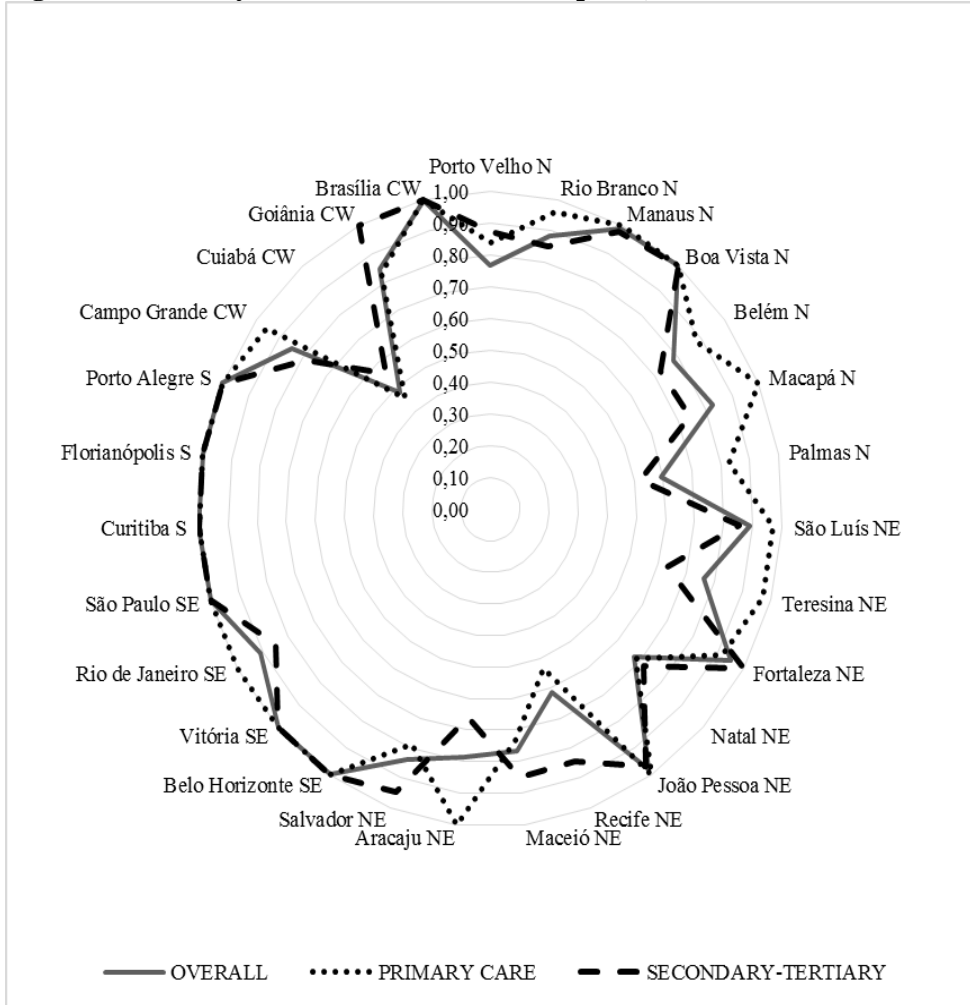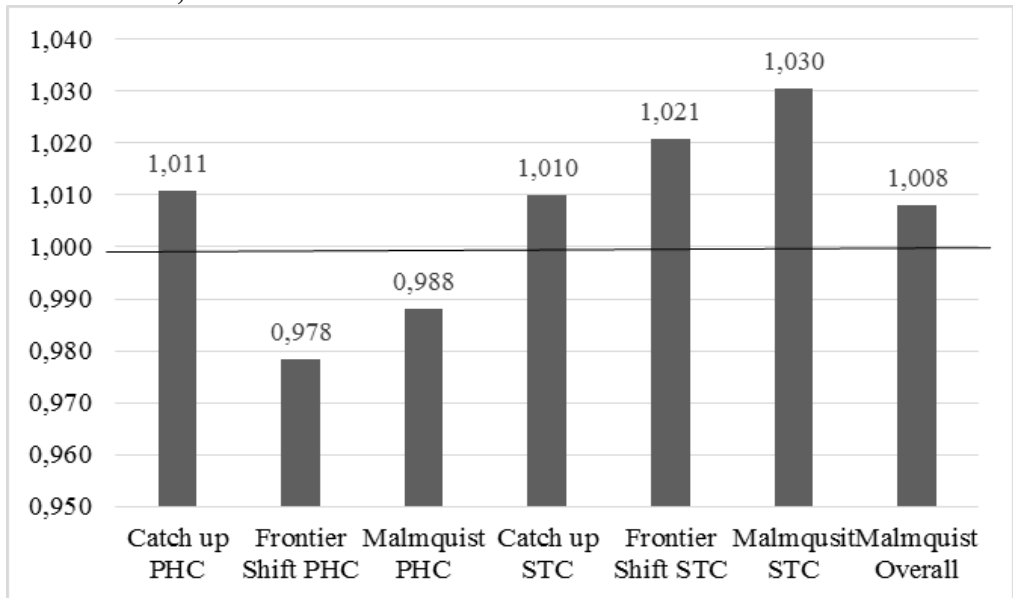

Source: Authors' own study

Figure 4. DEA efficiency scores - Brazilian state capitals, 2008-2013.

Source: Authors' own study

Figure 5. Malmquist Index, and its components, Brazilian capitals- 2008 to 2013: Overall, PHC and STC.
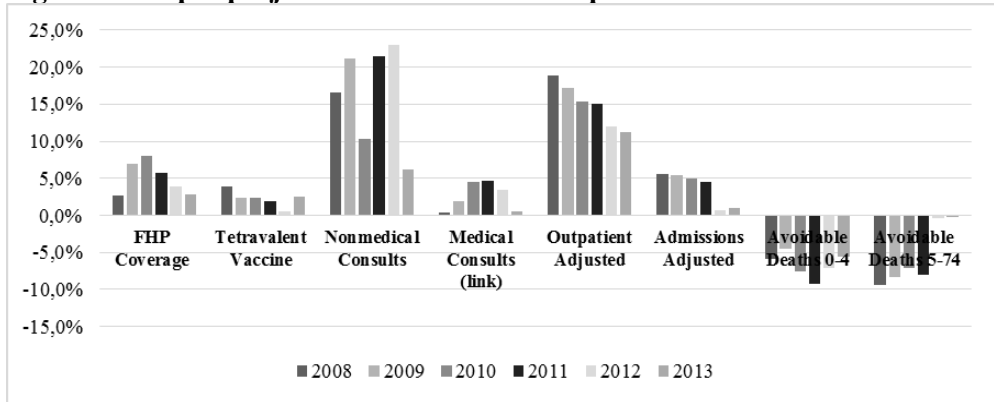

Source: Authors' own study

Figure 6 presents average projections to achieve the best production frontier across the years. For service delivery outputs, there is scope to a considerable large expansion in the number of nonmedical consultations (by 16.4%) and outpatients adjusted procedures (by 14.9%), which are PHC and STC outputs, respectively. The need to increase adjusted outpatient care is especially important because it is the main way of entrance into the STC level of care, by means of specialized consults and exams, suggesting there is a bottleneck at this point of the health network.

Note that medical consults, which have reduced their absolute numbers from 2008 to 2013 (Table 2) had to increase their production by only 2.5%, in average. As a link variable between PHC and STC, the projected values can be positive or negative. In order to achieve better efficiency scores, some capitals "asked" for reduction of the medical consults when there was overload of that variable to STC care, compatible with the hypothesis of a possible bottleneck in the interface between levels. Across the analyzed years, 7% to 22% of the capitals had negative projections, with magnitudes varying from to -3.3% to -65.5% of the actual values. The same capitals with negative projections had highest needs to increase the adjusted outpatient procedures (which are referenced from PHC physicians), corroborating the bottleneck effect hypothesis.

For both STC service delivery variables, there is a clear gradient towards decaying the projection needs in the analyzed period, compatible with the positive STC frontier shift shown above. For outcome quality health indicators, avoidable deaths still have to decrease by 5.9% and 9.4% for 0-4 years-old and 4-74 years-old, respectively, in 2013 projections.

Finally, in terms of expenditures (input slacks), there is scope to reduce PHC expenditures and STC expenditures by approximately 1.2 and 1.5 %, respectively, to produce the same level of outputs. These negative values for PHC and STC expenditures (input variables) consist of projections onto Pareto-inefficient portions of the frontier, which means that the same efficient scores would be found if the expenditures were reduced by the outlined percentage.

Figure 6: Output projections to reach the best practice frontier

Source: Authors' own study

Instead of average projections, individual DMU analysis can help to understand its behavior. When analyzing each capital in separate, the focus on projections will depend on the efficiency level and the pattern of inefficiency (either PHC or STC), facilitating establishment of goals to improve efficiency (Figure 7). For example, Rio de Janeiro, efficient in PHC, needed to increase outpatient procedures by 50% and adjusted admissions by 20% and reduce hospital avoidable deaths by 30%, in 2008. The gradual adjustment of these figures guaranteed STC efficiency in 2012 (Rio de Janeiro´s STC catch-up equals 1,071). On the other side, Fortaleza needed to increase PHC outputs, especially from 2008 to 2010 (HFP coverage by 30% and tetravalent vaccination, by 10%). Once accomplished the projected goal, Fortaleza appeared PHC efficient from 2010 until 2012, but did not maintain PHC efficiency in 2013 (the capital did not keep pace with the FHP coverage; so, Fortaleza´s PHC catch-up remained stationary = 1,001).

Figure 7. Overall, PHC and STC efficiency in Rio de Janeiro-SE and Fortaleza-NE, 2008-2013.
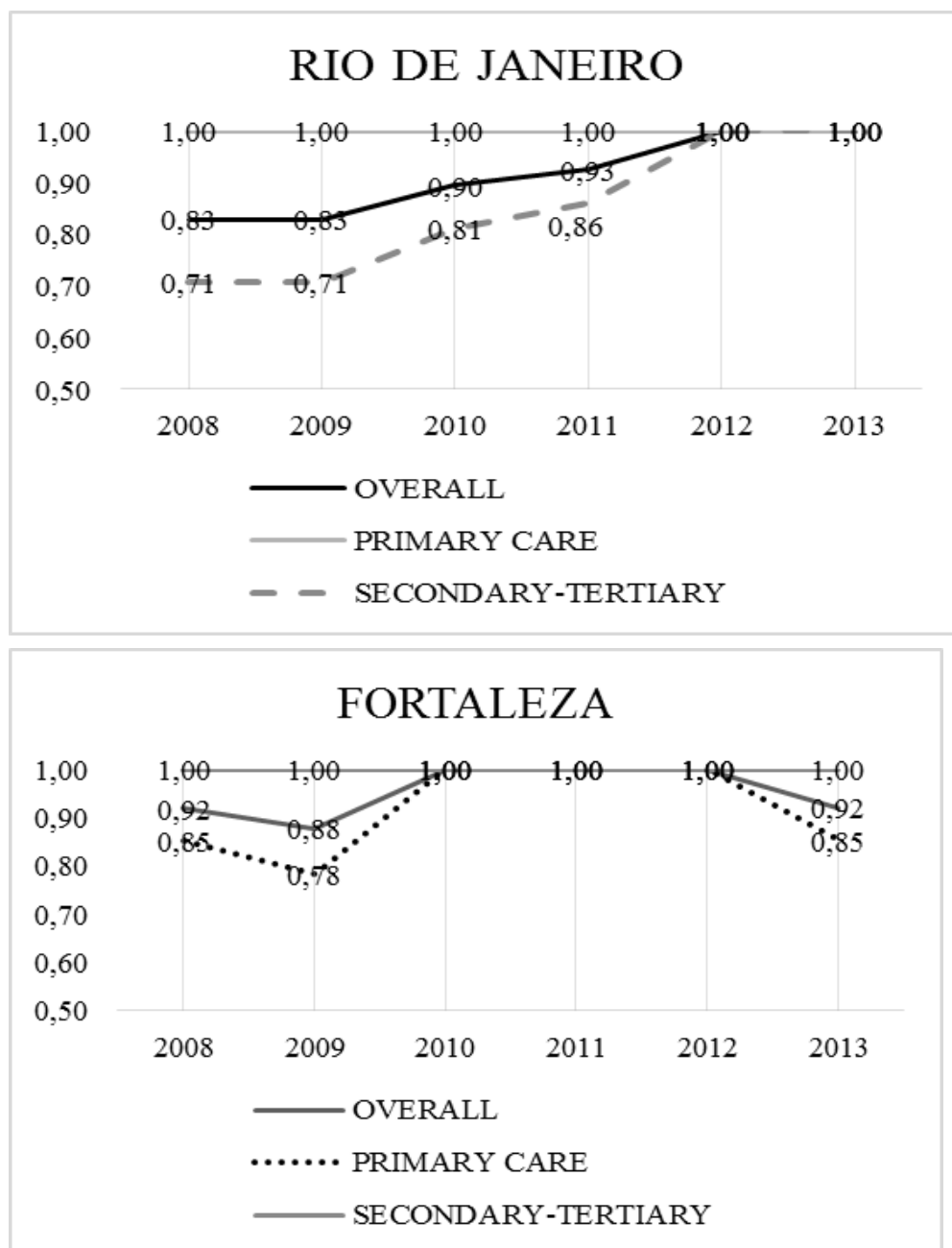
Source: Authors' own study

4. Discussion and final remarks

Health Care Systems can be thought as complex adaptative systems which do have many interacting agents and components in a changing environment (Peters 2014: 2). Resources (money, qualified personnel, facilities), biomedical knowledge (epidemiology, evidence-based medical practice) and population (health and sick) are inputs for the system; their interconnection determines capacity do learn an adapt to a sociocultural environment; and the latter exerts a powerful feedback effect on its inputs, dynamics and outputs (Costa 2012: 31). In this context, Data Envelopment Analysis is shown as a decision science method that carries few required model parameter choices while providing and effective means to analyze otherwise hard to model decision problems: it suits problems with a large number of input and output variables, heterogeneous data types within the model and differing scales while at the same time not being sensitive to distribution, autocorrelation and collinearity, issues found in parametric models (Alexander 2008: 109). While deepening interconnectivity approach, it is not surprising that a survey of the first forty years of scholarly literature has pointed out the use of network models as one of the main fields of current studies and future trends in DEA (Emrouznejad, Yang 2018: 4).

Most frequently, DEA health applications research published deal with cross-geographic (countries, states, municipalities), cross-facility (hospitals, primary care centers, nursing care, and so on) or health human resources (general practitioners, surgeons, specialists) comparisons, and use intermediate outputs (services delivered instead of health improvement), with little thoughts to model specifications (Hollingsworth 2008: 1108-1111). In Brazil, for public spending assessment, there are published DEA models to examine primary care efficiency (Varela et al. 2009) or outpatient productivity, either primary outpatient care or medium/high complexity (Ferreira, Pitta 2008). For secondary-tertiary performance analysis, hospitals are the most frequently used DMUs (Gonçalves et al. 2007; Cesconetto et al. 2008; Guerra et al. 2012; Souza et al. 2013; de Souza et al. 2016), although there are initiatives for specific high complexity procedures, as transplants (Costa et al., 2014). Network DEA model was used by Lobo et al. (2010, 2016) with university hospitals aiming to integrate teaching and health care dimensions and to aid financing decisions. As

each level or procedure was analyzed separately in each study, there would be limitations to infer the Brazilian health system as a whole from the sum of partial results of its components: primary care, outpatient, hospital, transplants, university hospitals.

In this study, DNSBM DEA model intended to represent Brazilian health care architecture in order to bring insights on how to monitor health systems achievements by linking PHC and STC levels of care through the SUS production function. The basic assumption was that the health system has its own components organized at a hierarchical structure, all interconnected and influenced by the others. The network model was a way to depict the relationships between levels of care, by opening the black box, instead of analyzing each level separately, which could compromise the interpretation of results. Additionally, the network DEA model allowed to consider, for different levels of care, intermediary or operational outputs (intermediary outputs or service delivery indicators) and quality outcomes (avoidable deaths).

During the analyzed period, there was huge investment on PHC and STC expending (above fifty percent increase), and the simultaneous evaluation of various outputs was necessary to conclude for or against efficiency in the use of resources. Once there are more resources, operational outputs changes can be observed in a short timeframe, but health indicators usually take more time – sometimes years or decades – to show steady variation clearly attributable to health policies undertaken.

Concerning operational outputs, at PHC level, there was a clear priority for FHP coverage in the analyzed period. It was noteworthy the decrease in PHC medical consults, problem that might be even bigger in poorer and smaller municipalities away from metropolitan areas. At STC level, both outputs – outpatients and admissions – increased substantially, and the need to augment decayed across time.

Concerning final outputs, avoidable deaths amelioration was evident for children under 5 years old. This result is attributed worldwide to high immunization coverage of the National Immunization Program, following the forth Millennium Development Goal (MDG) (Castro Lobo et al., 2014: 55), but there was projection to further decrease in 2013 (5.9%).

Given the 0.86 average score, there is still a great room to enhance efficiency. Individual analysis shows there was substantial differences among capitals, possibly correlated with socioeconomic disparities, given that inefficiency, especially for STC, prevails in poorer Northeast and North regions. Variations in time showed trends towards catching-up with the frontier at both, PHC and STC levels, but the technological change or positive frontier shift was only observed for STC level (which consumes about 90% of the budget) and at the richest regions. This may have resulted from the fact that managers invested in high complexity facilities inside capitals, which are STC reference for surrounding smaller municipalities and, not rarely, for all state municipalities.

The bottleneck effect seems to be an important challenge to overcome in Brazilian health system. When the entrance door to STC is blocked, individuals will enter the system directly from the emergency rooms, instead of being transferred from PHC level of care, distorting the system´s architecture as a whole. The reasons range from incapacity of PHC services to manage and coordinate patient care and tendency to focus on high-cost procedures (Paim et al. 2011: 1791).

The bottleneck hypothesis can be evaluated against other data. The Brazilian average number of physicians per 1,000 population equals 1.83. The same capitals that "asked" to reduce the number of PHC medical consults and to increase the outpatient procedures belong to states that have the lowest ratio of physicians per 1,000 population, varying from 0.58 (São Luis – NE) to 1.54 (Campo Grande – CW). In short, there is shortage of PHC physicians in these capitals and, simultaneously, overload of patients to access STC level of care; a scenario perfectly compatible with a bottleneck.

The observed input slacks (1.2% for PHC and 1.5% for STC) represent possible waste that could somewhat be invested in other healthcare activities.

For policy purposes, DNSBM DEA model can be used as an efficiency-based strategic tool for diagnosis and planning, to set and evaluate the accomplishment of service delivery goals, by comparing parameters of production among peers, vis-à-vis financing in one side (inputs for the production function), and health indicators on the other side (ultimate desired outputs). Assuming that each municipality sign annual management contracts with the public health manager, establishing budget

and quantitative volumes for service delivery, the study of projections for the goals under commitment is useful to understand and negotiate possibilities and benchmarks. The examples of Rio de Janeiro and Fortaleza showed how to monitor accomplishment of goals across time. Besides, policies to stimulate PHC medical consults in remote areas and intensify access and coordination between levels of care are important pathways to enhance efficiency.

For future research, new models can be designed according to diverse health systems architectures; diverse variables may be inserted according to policy priorities and management arrangements inside the system; and new boundaries must be defined (state, regions, microregions) to be compared. The important lesson that must always be kept in mind is that, no matter how much complex and uncertain the real world is, researchers need to understand it in depth and try to mimic subsets of it in order to have answers acceptable to managers and stakeholders (to guarantee face validity) instead of bringing the problem to academic silos and becoming distant from those who apply the solutions.

References

Afonso A.S.T., Aubyn M. (2005), Non-parametric approaches to education and health expenditure efficiency in OECD countries, „Journal of Applied Economics", vol. 8 no. 2, pp. 227-246.

Alexander M. (2008), Complex decision making using non-parametric data envelopment analysis, in: Complex decision making. Theory and practice, ed. Qudrat-Ullah H., Spector J.M., Davidsen P.I., Springer, Cambridgepp: 97-112.

Castro Lobo M.S., Estellita Lins M.P.E., Menegolla I.A. (2014), A new approach to assess the performance of the Brazilian National Immunization Program (NIP), „Socio-Economic Planning Sciences", vol. 48 no. 1, pp. 49-56.

Cesconetto A., Lapa J.D.S., Calvo M.C.M. (2008), Evaluation of productive efficiency in the Unified National Health System hospitals in the State of Santa Catarina, Brazil, „Cadernos de Saude Publica", vol. 24 no. 10, pp. 2407-2417.

Cooper W.W., Seiford L.M., Tone K. (2007), Data envelopment analysis. A comprehensive text with models, applications, references and DEA-Solver Software, 2. edition, Springer, New York.

Costa C.K.F., Balbinotto Neto G., Sampaio L.M.B. (2014), Efficiency of Brazilian states and the federal district in the public kidney transplant system based on DEA and the Malmquist index, „Cadernos de Saúde Pública", vol. 30 no. 8, pp. 1667-1679.

Costa J. (2012), Systems pathology. A critical review. „Molecular Oncology", vol. 6, pp. 27-32.

DATASUS – Ministry of Health / Brazil Morbidity Indicators Data System. http://www2.datasus.gov.br/DATASUS/index.php?area= [6.12.2017].

De Souza P.C., Scatena J.H.G., Kehrig R.T. (2016), Data envelopment analysis application to evaluate the efficiency of SUS's hospitals in the state of Mato Grosso, Brazil, „Physis", vol. 26 no. 1, pp. 289-308.

Emrouznejad A., Yang G. (2018), A survey and analysis of the first 40 years of scholarly literature in DEA: 1978-2016, „Socio-Economic Planning Sciences", vol. 61 no. 1, pp. 1-5.

Emrouznejad A., De Witte K. (2010), COOPER-framework. A unified process for non-parametric projects, „European Journal of Operational Research", vol. 207 no. 3, pp. 1573-1586.

Emrouznejad A., Amin G.R. (2009), DEA models for ratio data. Convexity consideration, „Applied Mathematical Modelling", vol. 33 no. 1, pp. 486-498.

Färe R., Grosskopf S. (2000), Network DEA, „Socio-Economic Planning Sciences", vol. 34, pp. 35-49.

Färe R, Grosskopf S., Lindgren B., Roos P. (1992), Productivity changes in Swedish pharmacies 1980-1989. A non-parametric Malmquist approach, „Journal of Productivity Analysis", vol. 3, pp. 85-101.

Färe R., Grosskopf S., Norris M., Zhang Z. (1994), Productivity growth, technical progress, and efficiency change in industrialized countries, „American Economic Review", vol. 84, pp. 66-83.

Ferreira M.P., Pitta M.T. (2008), Evaluation of the technical efficiency of the outpatient services of SUS in São Paulo, „São Paulo em Perspectiva", vol. 22 no. 8, pp. 55-71.

Gonçalves A.C., Noronha C.P., Lins M.P.E., Almeida R.M.V.R. (2007), Data envelopment analysis for evaluating public hospitals in Brazilian state capitals, „Revista de Saude Publica", vol. 41 no. 3, pp. 427-435.

Gragnolati M., Lindelow M., Couttolenc B. (2013), 20 years of health system reform in Brazil. An Assessment of the Sistema Único de Saúde, The World Bank, Washington, DC, http://documents.worldbank.org/curated/pt/909701468020377135/Twenty-years-of-health-system-reform-in-Brazil-an-assessment-of-the-sistema-unico-de-saude [6.12.2017].

Guerra M., de Souza A.A., Moreira D.R. (2012), Performance analysis. A study using data envelopment analysis in 26 Brazilian hospitals, „Journal of Health Care Finance", vol. 38 no. 4, pp. 19-35.

Hollingsworth B. (2008), The measurement of efficiency and productivity of health care delivery, „Health Economics", vol. 17, pp. 1107-1128.

Lobo M.S.C., Lins M.P.E., Silva A.C.M., Fiszman R. (2010), Assessment of teaching-health care integration and performance in university hospitals, „Revista de Saúde Pública", vol. 44 no. 8, pp. 581-590.

Lobo M.S.C, Rodrigues H.C, Gazzola E.C.A., de Azeredo J.A, Lins M.P.E. (2016), Dynamic network data envelopment analysis for university hospitals evaluation, „Revista de Saúde Pública", vol. 50 no. 22, pp. 1-11.

Lobo M.S.C., Lins M.P.E. (2010), Epistemic dialog between health services and operations research, „Pesquisa Operacional", vol. 30 no. 2, pp. 371-386.

Malta D.C., Santos M.A.S., Stopa S.R., Vieira J.E.B., Melo E.A., Reis A.A.C. (2016), A Cobertura da Estratégia de Saúde da Família (ESF) no Brasil, segundo a Pesquisa Nacional de Saúde, 2013 (Family Health Strategy Coverage in Brazil, according to the National Health Survey, 2013), „Ciência & Saúde Coletiva", vol. 21 no. 2, pp. 327-338.

Malta D.C., Sardinha L.M.V., Moura L., Lansky S. et al. (2010), Atualização da lista de causas de mortes evitáveis por intervenções do Sistema Único de Saúde (Update of avoidable causes of deaths due to interventions at the Brazilian Health System), „Epidemiologia e Serviços de Saúde", vol. 19 no. 2, pp. 173-176.

OECD (2015), OECD economic surveys. Brazil 2015, OECD Publishing, Paris, http://dx.doi.org/10.1787/eco_surveys-bra-2015-en [6.12.2017].

Paim J., Travassos C., Almeida C., Bahia L., Machado J. (2011), The Brazilian health system. History, advances and challenges, „The Lancet", vol. 377, pp. 1778-1797.

Peters D.H. (2014), The application of systems thinking in health. Why use systems thinking?, „Health Research Policy and Systems", vol. 12 no. 51, pp. 2-6.

Souza F.V., de Melo M.D., Araújo A., da Silva M. (2013), Efficiency of public spending on hospital care. A study in Brazilian state capitals in the period 2008 to 2010, „Holos", vol. 1, pp. 203-216.

Tone K. (2001), A slacks-based measure of efficiency in data envelopment analysis, „European Journal of Operational Research", vol. 130, pp. 498-509.

Tone K., Tsutsui M. (2009), Network DEA. A slacks-based measure approach, „European Journal of Operational Research", vol. 197, pp. 243-252.

Varela P.S., Martins G.A., Fávero L.P.L. (2009), Production efficiency and financing of public health. An Analysis of small municipalities in the state of São Paulo – Brazil, „Health Care Management Science", vol. 13 no. 2, pp. 112-123.

EFFICIENCY ANALYSIS OF PUBLIC HEALTH SPENDING IN BRAZILIAN CAPITALS

Analiza efektywności wydatków na służbę zdrowia w brazylijskich stolicach stanowych w oparciu o metodę Network Data Envelopment Analysis

Streszczenie

Cel: W 1988 roku Brazylia wdrożyła dogłębne zmiany w organizacji oraz finansowaniu systemu służby zdrowia, powołując do życia Zunifikowany System Opieki Zdrowotnej („Sistema Unico de Saúde" – SUS), zakładający powszechny dostęp do służby zdrowia. Stopniowemu rozszerzaniu systemu opieki zdrowotnej oraz uprawnień do usług medycznych towarzyszyła debata dotycząca odpowiedniego poziomu wydatków rządowych i wydajności systemu.

Metodyka badań: W badaniu wykorzystano zmienne efekty skali zorientowane na wyniki, model Slacks-Based-Measure oparty na dynamicznej sieciowej metodzie obwiedni danych (ang.: Data Envelopment Analysis (DEA)) dla danych z lat 2008-2013, aby zobrazować zależność, jaka zachodzi pomiędzy różnymi poziomami opieki (podstawowa opieka zdrowotna (ang. primary health care (PHC)) oraz opieka zdrowotna drugiego i trzeciego stopnia (ang. secondary-tertiary health care (STC)). Jednostkami podejmującymi decyzje są brazylijskie stolice stanowe, które wdrażają kluczowe założenia polityki zdrowotnej oraz wspierają pacjentów z okolicznych, mniejszych jednostek administracyjnych, szczególnie w zakresie STC. Nakłady stanowią budżety PHC i STC, natomiast wynikami są wynikające z nich usługi oraz przypadki zagrożenia życia, w których udało się uratować pacjentów. Powiązana zmienna to konsultacje medyczne w ramach PHC, drzwi wejściowe do systemu oraz strażnik bramy do bardziej kompleksowych poziomów opieki. Dynamiczny model pozwala oceniać wydajność w czasie.

Wnioski: Ogólny stan wyniósł 0,86, przy czym dla PHC kształtował się na poziomie 0,90, a dla STC 0,85 (SD=0,15). 8 z 27 stolic okazało się w pełni wydajnych. Stolice zdołały zwiększyć wyniki w obu poziomach opieki zdrowotnej, ale tylko STC doświadczyło pozytywnej zmiany technologicznej (frontier shift > 1). Zmienna powiązana wykazała wąskie gardło pomiędzy poziomami opieki. Projekcje dotyczące granic (frontier shift) pozwoliły ustalić własną diagnozę dotyczącą zarządzania oraz cele związane z finansowaniem i rozwojem.

Wartość artykułu: Modele sieciowe naśladują hierarchicznie zorganizowane systemy opieki zdrowotnej. Wykorzystanie wyników wspiera politykę służby zdrowia.

Słowa kluczowe: Network DEA, pomiar wydajności, rozwój zdrowotny i gospodarczy, służba zdrowia
JEL: H21, H51, O10, O54