

Desempenho dos municípios paulistas: uma avaliação de eficiência da atenção básica à saúde

Patrícia Siqueira Varela
Gilberto de Andrade Martins
Luiz Paulo Lopes Fávero

Recebido em 15/fevereiro/2011
Aprovado em 26/março/2012

Sistema de Avaliação: *Double Blind Review*
Editor Científico: Nicolau Reinhard

DOI: 10.5700/rausp1063

A autora Patrícia Siqueira Varela agradece o apoio institucional recebido durante o pós-doutorado realizado no Programa de Mestrado Acadêmico e Doutorado em Administração Pública e Governo da Fundação Getúlio Vargas de São Paulo.

RESUMO

Neste trabalho, teve-se por objetivo construir uma metodologia de avaliação de desempenho dos municípios paulistas quanto à eficiência técnica na aplicação de recursos públicos nas ações de atenção básica à saúde e analisar a influência de variáveis não controláveis no processo de produção em tal área. A eficiência técnica é um dos parâmetros de avaliação de desempenho dos gestores públicos, refletindo a capacidade de uma entidade obter máximos *outputs* com o menor consumo de *inputs*. O alcance de tal métrica pode ser prejudicado ou favorecido pelas variáveis ambientais ou não controláveis que, se não forem consideradas na avaliação de desempenho, podem gerar vieses. Nesse sentido, por meio da metodologia *Data Envelopment Analysis* (DEA) em dois estágios, os escores de eficiência dos municípios foram estimados e depois ajustados com o uso da análise de regressão. Os resultados indicaram que seria possível aumentar, consideravelmente, a quantidade de serviços prestados à população sem a necessidade de novas dotações orçamentárias na maioria dos municípios. Além disso, verificou-se que a maior proporção de idosos em uma jurisdição torna a prestação de serviços mais cara; por sua vez, maiores densidade populacional, grau de urbanização e escala dos estabelecimentos de saúde favorecem o gasto público com eficiência. Os cinco municípios paulistas considerados mais eficientes foram Tuiuti, Nova Guataporanga, Sabino, Lins e Santos.

Palavras-chave: controladoria, avaliação de desempenho, eficiência, atenção básica à saúde, municípios.

1. INTRODUÇÃO

O Sistema Único de Saúde (SUS) é caracterizado por um sistema complexo de relações intergovernamentais. O governo federal estabelece normas e age como coordenador das ações dos entes federativos, buscando a concretização

Patrícia Siqueira Varela, Professora Adjunta da Universidade Federal de São Paulo – Campus Osasco (CEP 06110-295 – Osasco/SP, Brasil), com Mestrado e Doutorado em Ciências Contábeis pela Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo e Pós-Doutorado em Administração Pública e Governo pela Fundação Getúlio Vargas, é Professora Orientadora do Programa de Pós-graduação em Gestão e Informática em Saúde da Universidade Federal de São Paulo.

E-mail: psvarela@unifesp.br

Endereço:

Universidade Federal de São Paulo

Campus Osasco

Rua Angélica, 100

06110-295 – Osasco – SP

Gilberto de Andrade Martins é Professor Titular do Departamento de Contabilidade e Atuária da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (CEP 05508-010 – São Paulo/SP, Brasil), Professor do Programa de Pós-Graduação em Ciências Contábeis.

E-mail: martins@usp.br

Luiz Paulo Lopes Fávero, Professor Livre-Docente da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA-USP) (CEP 05508-010 – São Paulo/SP, Brasil), com Mestrado e Doutorado em Administração pela FEA-USP e Pós-Doutorado em Econometria Financeira pela *Columbia University*, é Professor do Programa de Pós-Graduação em Controladoria e Contabilidade da FEA-USP.

E-mail: lpfaver@usp.br

do atendimento universal, equitativo e integral ao cidadão. Os municípios são os principais provedores de serviços na área da saúde, com responsabilidades exclusivas quanto à subfunção atenção básica à saúde.

Como política setorial de governo, o SUS tem sido influenciado pelos esforços recentes de modernização da administração pública, tendo como um de seus pilares a disseminação dos princípios da administração pública gerencial, cujos focos são a gestão voltada para resultados e o uso eficiente dos recursos. Nesse contexto, o programa aparece como o elemento integrador dos planos e orçamentos e como unidade de gestão das ações governamentais. Assim, a avaliação de desempenho quanto a efetividade, eficácia e eficiência dos programas governamentais ganha destaque.

Todavia, a determinação de parâmetros de avaliação pela controladoria na gestão pública não é uma tarefa simples, pois o tipo de controle aplicável às atividades do setor público depende de quatro critérios: ambiguidade dos objetivos, mensuração dos *outputs*, conhecimento dos efeitos das intervenções e repetição das atividades.

A avaliação de desempenho envolve dificuldades relacionadas à mensuração dos produtos diretos e, mormente, os finais dos programas governamentais, à avaliação da qualidade e ao relacionamento entre recursos orçamentários e efetividade. No caso específico da saúde, a determinação dos centros de *inputs-outputs* e sobre o que cada gestor público pode ser responsabilizado também se torna um problema, dado que existem interações entre os diversos sistemas de saúde (municipais, estaduais e federal), com a existência de fluxos de pacientes e recursos entre eles, de acordo com a capacidade de cada um de prestar serviços de média ou alta complexidades.

Para contornar essa última situação, optou-se por abordar a subfunção atenção básica à saúde, que é exclusiva de cada município e, assim, os gestores públicos podem ser responsabilizados pelo que acontece em seu território, desde que seja fruto de seu poder discricionário. Além disso, o foco foi nos *outputs* diretos da produção em saúde, simplificando muitas das restrições ao processo de avaliação de desempenho e *accountability*, principalmente quanto à defasagem entre as ações e os seus resultados efetivos.

O orçamento é considerado um dos principais veículos do processo de planejamento, execução e controle. Todavia, a determinação de metas depende da conciliação entre os aspectos motivacionais e o planejamento realístico dos números do orçamento. São sugeridos dois princípios para resolver esse problema: remover a ênfase em metas fixas e presentes e substituí-las por padrões de desempenho por *benchmark* e ajustar as avaliações de desempenho para refletir a experiência de operação vigente e circunstâncias econômicas enfrentadas durante o período.

Nesse sentido, o objetivo neste estudo foi: construir uma metodologia de avaliação de desempenho dos municípios paulistas quanto à eficiência técnica na aplicação de recursos

públicos nas ações de atenção básica à saúde e analisar a influência de variáveis não controláveis no processo de produção em tal área.

A avaliação comparativa de desempenho permite identificar os pontos fracos da gestão, assim como aqueles que servem como referência para outras jurisdições, concretizando uma das vantagens das estruturas federalistas, como é o caso do Brasil, o aprendizado pela diversidade de experiências. Neste trabalho, inova-se ao buscar usar o modelo DEA (*Data Envelopment Analysis*) em dois estágios para verificar a influência de variáveis não controláveis na produção de *outputs* diretos das ações da área da saúde, por exemplo, o percentual de população rural. Isso porque tais variáveis podem alterar as condições de produção do serviço público de saúde e, se não forem consideradas, a análise dos índices de eficiência pode apresentar viés.

Uma das principais dificuldades para aplicar a técnica é a seleção de variáveis representativas do modelo de produção em atenção básica à saúde dos municípios, ou seja, os recursos consumidos (*inputs*) e os produtos gerados (*outputs*) pelas entidades. Sendo assim, a metodologia sugerida ganha relevância pelas variáveis consideradas no modelo DEA, as quais foram escolhidas com base na teoria econômica sobre função de produção (base de avaliação dos programas sociais) e análise de eficiência, além de considerar as especificidades do sistema de saúde brasileiro. Adicionalmente, tratar de uma área tão complexa quanto a saúde e buscar avaliar a eficiência dos municípios lança luz sobre as possibilidades de transposição das abordagens tradicionais da controladoria, desenvolvidas para empresas, para organizações governamentais.

Espera-se que este artigo possa contribuir para a discussão sobre bases teórico-empíricas para a avaliação de desempenho no âmbito do setor público, o que pode favorecer o aumento do controle social sobre os gastos públicos, a melhoria dos serviços prestados e a redução de custos. Além disso, a metodologia, ora apresentada, pode ser uma alternativa para as entidades de controle externo, mormente tribunais de contas, realizarem auditoria operacional nos entes públicos sob sua jurisdição.

2. AVALIAÇÃO DE DESEMPENHO DE ENTIDADES DO SETOR PÚBLICO

A controladoria é de difícil definição, podendo ser visualizada sob dois enfoques: como um órgão administrativo ou como uma área de conhecimento humano. Borinelli (2006, p.105), após uma análise crítica das definições apresentadas nas literaturas brasileira e internacional, afirma que a

“controladoria é um conjunto de conhecimentos que se constituem em bases teóricas e conceituais de ordens operacional, econômica, financeira e patrimonial, relativas ao controle do processo de gestão organizacional”.

Chenhall (2007, p.164) menciona o uso, algumas vezes intercambiável, dos termos *Management Accounting* (MA), *Management Accounting Systems* (MAS), *Management Control Systems* (MCS) e *Organizational Controls* (OC). Para ele, MCS é um termo mais amplo, que abarca MAS e também inclui outros controles, tais como de pessoas ou grupos. Além disso, a definição de MCS

“evoluiu para além daquela de simplesmente focar na provisão, mais formal, de informação quantificável financeiramente com o intuito de auxiliar a tomada de decisão gerencial para outra que compreende um escopo muito mais amplo de informação” (CHENHALL, 2007, p.165, tradução nossa).

Algumas atividades de controladoria são realizadas por várias ou todas as unidades organizacionais, na forma de processo, e outras podem ser exercidas por uma área formalmente constituída em uma organização, principalmente quando se trabalha em termos macro. Também a área de controladoria é responsável por consolidar as atividades realizadas de forma descentralizada (BORINELLI, 2006). No setor público, por imposições constitucionais e legais, existem órgãos de controle, tanto interno quanto externo, que precisam lidar com o grande desafio de avaliar o desempenho daqueles que respondem pela aplicação dos recursos públicos, não apenas quanto à legalidade do gasto, mas principalmente em relação à sua qualidade. Os órgãos públicos de controle brasileiros têm vivido um processo de aprendizagem sobre a avaliação de desempenho no âmbito das diferentes políticas públicas em relação aos benefícios gerados à população e seus respectivos custos, o que não é possível fazer com o uso de abordagens tradicionais de controle gerencial para as empresas.

Conforme afirma Hofstede (1981, p.193), tais abordagens são, usualmente, insuficientes para muitas das atividades desenvolvidas por organizações públicas e não lucrativas. O grau de complexidade do tipo de controle a ser exercido depende de quatro critérios: não ambiguidade dos objetivos, mensuração dos *outputs*, conhecimento do efeito das intervenções e repetição das atividades. Quando um ou mais desses critérios não são satisfeitos, são exigidos sistemas mais complexos de controle.

Os objetivos podem ser ambíguos por causa dos conflitos de valores ou interesses entre aqueles que opinam sobre a atividade, da falta de conhecimento sobre como a atividade pode levar a um fim desejado e das rápidas mudanças ambientais que forcem o estabelecimento de novos objetivos ou tornam os existentes obsoletos (HOFSTEDÉ, 1981, p.194).

Os *outputs* são mensuráveis quando se pode quantificá-los. Entretanto, muitas atividades, tanto em organizações públicas quanto privadas, têm *outputs* que só podem ser definidos qualitativamente e em termos vagos. Nesses casos, somente os *inputs* podem ser mensurados. Por exemplo, a efetividade

dos programas quando não se pode separar os impactos de cada um deles sobre o bem-estar da sociedade (HOFSTEDÉ, 1981, p.195).

Em relação ao conhecimento do efeito das intervenções, Hofstede (1981, p.195) salienta que o controle gerencial pressupõe que os esforços alocados para uma atividade podem ser redirecionados para outra se as metas não são alcançadas. Contudo, isso implica que o gestor conheça como a intervenção produz a correção desejada, o que implica entender a tecnologia da atividade.

Por fim, a repetição periódica das atividades possibilita um efeito aprendizagem, o que se estende para mudança de configuração de elementos repetitivos, tais como em muitas atividades profissionais, por exemplo, medicina e odontologia. O orçamento corrente é tipicamente um processo repetitivo (HOFSTEDÉ, 1981, p.195).

O SUS usa diversos mecanismos para contornar o problema da ambiguidade de objetivos, tais como determinação de diretrizes pela União, realização de pactos entre os entes federados, incentivos à implementação de alguns tipos de estratégia de atenção e a existência dos conselhos de saúde, os quais são formados por profissionais, gestores e cidadãos.

Supondo a não ambiguidade dos objetivos, mesmo assim a organização das ações em programas de saúde não é uma tarefa fácil. É preciso identificar os produtos de cada ação e como e quanto esses produtos intermediários, em conjunto, irão contribuir para gerar modificações no estado de saúde da população, devendo ser analisadas as interferências mútuas entre os programas.

Hofstede (1981, p.202) ressalta o papel do orçamento como principal veículo de controle gerencial e o distingue em três categorias: investimentos, operacional para centros de *input-output* e operacional para centros de *input*. Essas duas últimas também denominadas de centros de lucros ou de despesas, respectivamente. As restrições quanto à mensuração dos *outputs* das atividades realizadas por entidades do setor público dificultam a definição de centros de *inputs-outputs*, fazendo com que o controle orçamentário tenha seu foco apenas no gasto, desconsiderando as realizações com a execução da despesa. Todavia, o movimento de transformação da administração pública burocrática para a gerencial enseja esforços no sentido de identificar o que se faz com a aplicação de recursos públicos, ou seja, que se identifiquem os produtos das ações governamentais e se trabalhe com centros de *inputs-outputs*.

Na nova abordagem da gestão pública, os programas são tratados como unidades de gestão, configurando-se como um centro de *input-output*. Contudo, outros tipos de centros de *input-output* ou *input* – unidade de saúde, municípios, regiões de saúde, estados, União – podem ser relevantes, o que vai depender do tipo de análise que se pretende fazer e do objeto de análise. Por exemplo, se o objetivo é controlar o limite legal de gastos municipais com a função saúde, os governos municipais são considerados centros de *inputs* na análise de desempenho.

Um centro de *input-output* ou *input* configura-se, na verdade, em um centro de responsabilidade que, de acordo com Anthony e Govindarajan (2008, p.129), “existe para cumprir um ou mais propósitos, tidos como seus *objetivos*”. Neste trabalho, o foco é avaliar os gestores públicos municipais como responsáveis pelo atendimento básico de saúde à população de sua jurisdição.

O orçamento público é um dos principais instrumentos do processo de gestão: planejamento, execução e controle. Assim como nos demais países, o orçamento público no Brasil orientou-se, inicialmente, pela necessidade de possibilitar aos órgãos de representação um controle político sobre o Executivo e a preocupação estava em manter o equilíbrio financeiro e evitar ao máximo a expansão dos gastos (GIACOMONI, 2000, p.64). Com a mudança de ênfase do orçamento público para planejamento e gerência, a avaliação de desempenho ganha destaque no contexto de aperfeiçoamento dos gastos públicos quanto à eficiência, à eficácia e à efetividade dos programas governamentais. De posse dos planos e orçamentos, dos parâmetros da avaliação e da mensuração dos resultados alcançados, é possível criar um juízo de valor e fazer julgamentos sobre as ações dos gestores em seus centros de responsabilidade e em determinado período de tempo.

Verbeeten (2008, p.430) elenca alguns propósitos das práticas de administração de desempenho: a definição de missões, objetivos e metas claros ajuda cada funcionário a entender o que a organização deseja e produz foco nas operações; os administradores públicos e políticos deveriam ser capazes de mostrar a finalidade do uso do dinheiro ao público (transparência / *accountability*) pela mensuração do desempenho no que diz respeito aos objetivos e metas; as organizações do setor público podem usar a mensuração do desempenho no processo de aprendizagem, a partir da identificação dos excessos da organização e dos aperfeiçoamentos necessários; e o sistema de avaliação de desempenho pode fornecer as bases para remuneração dos funcionários públicos.

Merchant e Otley (2007, p.791) afirmam que os sistemas de controle orientados para *accountability* buscam, antes de tudo, manter os indivíduos ou grupos de indivíduos responsáveis tanto por suas ações quanto por seus resultados ou aqueles que sua organização produz, sendo recompensados ou punidos de acordo com o desempenho da organização.

Merchant e Otley (2007, p.792) listam sete dos principais temas encontrados na literatura relativa à *accountability* na forma das seguintes perguntas:

- O que faz uma medida de desempenho boa (ou conjunto de medidas)?
- Por que os administradores geralmente são mantidos responsáveis por muito mais do que eles podem controlar?
- Metas de desempenho são necessárias e, se sim, o que as torna boas?
- O que se conhece sobre a escolha dos estilos de *accountability*?

- Quais são as questões-chave do desenho do sistema de incentivo e o que se conhece sobre elas?
- Como e por que os sistemas de controle diferem em diferentes cenários?
- Como se pode reconhecer progresso, isto é, inovações positivas como distintas de modas passageiras?

As respostas às três primeiras questões trazem alguns pontos de reflexão importantes para este trabalho e é sobre elas que se pretende discutir.

Em relação à qualidade das medidas de desempenho, Merchant e Otley (2007, p.792) mencionam que parece existir uma concordância geral de que boas medidas devem refletir o progresso em direção ao alcance dos objetivos da organização. A congruência da medida usada para fins de controle poderia ser julgada por sua correlação com o desempenho real da organização.

Ao tratar de organizações sem fins lucrativos, Merchant e Otley (2007, p.793) afirmam que o objetivo dessas entidades é cumprir uma missão e, assim, o desempenho financeiro é uma restrição. A mensuração dos *inputs* do processo é relativamente fácil, mas a mensuração do desempenho é mais difícil, mesmo que exista um acordo entre os eleitores quanto aos objetivos da organização. Isso porque o espaço de tempo entre a ação e os resultados é, algumas vezes, muito longo e a qualidade dos serviços fornecidos é de difícil mensuração.

Quanto à responsabilidade dos gestores, o mais antigo princípio de controle consiste em as pessoas responderem somente pelo que elas podem controlar. Antes do período de mensuração, as organizações desenham estruturas de responsabilidades para alinhar medidas de desempenho com os níveis de autoridade dos gestores (MERCHANT e OTLEY, 2007, p.793).

Na prática, entretanto, observa-se que muitos, talvez a maioria, dos gestores são considerados responsáveis por fatores que não podem controlar. As duas principais causas são, em primeiro lugar, a difícil separação entre efeitos controláveis e não controláveis sobre a medida de desempenho e, em segundo, o desejo das organizações de que os gestores respondam às mudanças nos fatores não controláveis, ou seja, influenciem a medida de desempenho, mesmo que o fator seja não controlável. Se os gestores podem, materialmente, influenciar os efeitos de um fator sobre o desempenho, deveriam ser considerados responsáveis pelos efeitos daquele fator, independentemente de ser ou não controlável (MERCHANT e OTLEY, 2007, p.794).

Por fim, as reflexões sobre o estabelecimento de metas de desempenho passam pela discussão de como combinar os aspectos motivacionais das metas com os números do planejamento realístico do orçamento. Fraser e Hope (2003, *apud* MERCHANT e OTLEY, 2007, p.794) argumentam que muitos dos problemas experimentados pelas organizações, na tentativa de gerenciar seus procedimentos de controle orçamentário, originam-se do fato de o orçamento representar uma meta **fixada**

(presente e imutável) em um mundo de mudança constante e incerteza. A solução para muitos desses problemas seria adotar alguns novos princípios de controle. O primeiro deles seria remover a ênfase em metas fixas e presentes e substituí-las por padrões de desempenho por *benchmark*, o que permitiria a implementação do segundo, defensavelmente mais importante, princípio. Isto é, as avaliações de desempenho poderiam ser ajustadas para refletir a experiência de operação vigente e circunstâncias econômicas enfrentadas durante o período.

Como visto até agora, a avaliação de desempenho e *accountability* do setor público, especificamente da área da saúde, envolve conhecer seus objetivos, a forma de funcionamento, a sistematização das informações em planos e orçamentos, as medidas de desempenho e os parâmetros da avaliação de desempenho. Neste trabalho, o foco está em avaliar a eficiência dos municípios paulistas na subfunção atenção básica, o que pressupõe conhecer o processo de produção de bens públicos.

3. PRODUÇÃO DE BENS PÚBLICOS

De acordo com Cohen e Franco (2004, p.77), os conceitos utilizados em avaliação de projetos sociais, e aí se incluem os serviços de saúde, foram desenvolvidos pela economia e estão associados à função de produção que vincula recursos, insumos, processos e produtos.

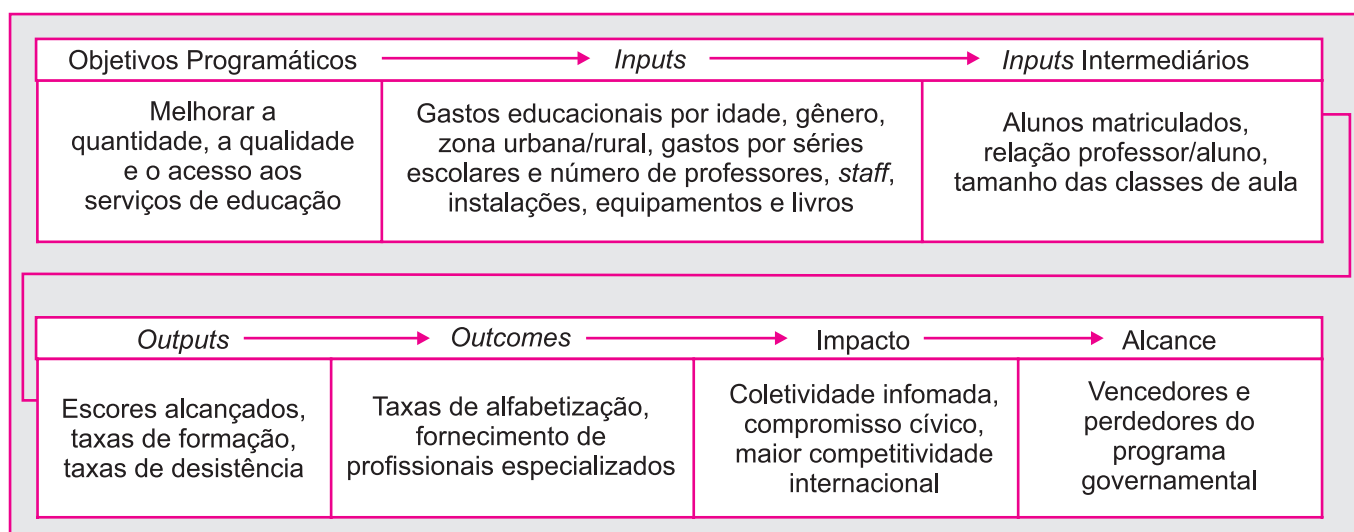
Shah (2007, p.11), ao tratar de transferências intergovernamentais vinculadas à obtenção de resultados, menciona que o administrador público de um governo receptor poderia examinar os diversos elementos de um programa a fim de determinar se os resultados seriam ou não alcançados, quais sejam:

- *inputs* (recursos usados para produzir *outputs*), incluindo *inputs* intermediários;
- *outputs* (quantidade e qualidade dos bens e serviços públicos produzidos e o acesso a eles);
- *outcomes* (consequências de médio/longo prazo para os contribuintes/consumidores da provisão dos serviços públicos ou grau de alcance dos objetivos do programa);
- impacto (objetivos do programa ou consequências de longo prazo);
- alcance (pessoas beneficiadas ou prejudicadas pelo programa).

Um exemplo da área de educação pode ser visualizado pela figura abaixo.

Cohen e Franco (2004, p.93), ao tratarem das saídas do processo de produção, destacam a distinção entre produtos, efeitos e impactos. Os efeitos são resultados da utilização dos produtos e o impacto, a consequência dos efeitos de um projeto, ou seja, o grau de consecução dos objetivos (estado desejado) em relação à população-meta.

Pelas abordagens de Cohen e Franco (2004) e Shah (2007), é possível perceber que, no processo de provisão de bens e serviços públicos, a dificuldade em mensurar os *outputs* surge de forma mais aguda do que na produção de bens privados. Bradford, Malt e Oates (1969) já haviam se preocupado com essa questão e buscaram esclarecê-la com o desenvolvimento de um modelo em que os *outputs* são separados entre diretos (*Output-D*) e de interesse final do cidadão (*Output-C*). Por exemplo, o produto final de um programa sentido por uma população não é o número de vacinas aplicadas, mas a redução da mortalidade infantil.



Aplicação da Cadeia Baseada em Resultados para a Educação

Fonte: Shah (2007, p.11).

Combinações diferentes de *inputs* (*Input-I*) – trabalho, capital, equipamentos etc. – geram vetores de bens e serviços diretamente produzidos (*Output-D*), por exemplo, consultas médicas, exames laboratoriais e visitas dos agentes comunitários de saúde. Esses produtos diretos do sistema de saúde, em conjunto com certas variáveis ambientais ou exógenas, determinam o *output* final (*Output-C*) sentido pela população, por exemplo, na redução da incidência de doenças cardiovasculares em um ambiente marcado por hábitos alimentares não salubres.

Com base no exposto anteriormente, as preferências dos indivíduos podem ser expressas como $U = U(C, Z)$, em que o vetor Z representa o nível de provisão de outros bens públicos, além do *Output-C*, e da quantidade de bens privados consumidos pelo indivíduo. O *Output-C*, por sua vez, pode ser expresso como $C_k = f_k(D, E)$, com E representando um vetor de variáveis ambientais e f_k uma função indicativa do grau do *output* final sentido pelo indivíduo ao usufruir os *Outputs-D* produzidos no ambiente E (BRADFORD, MALT e OATES, 1969, p.187).

Conforme Bradford, Malt e Oates (1969, p.187), em termos dessas definições, faz sentido multiplicar o vetor D (bens e serviços diretamente produzidos) por algum escalar, mas não o vetor C (bens e serviços de interesse final do cidadão). Por exemplo, para prover o dobro de consultas médicas, podem ser necessárias duas vezes o número de médicos. Se o dobro de D exigir menos que o dobro de *inputs*, existirá economia de escala na provisão de *outputs* diretos. Todavia, não dá para falar no dobro de C , já que isso depende de outras variáveis ambientais, mas é possível examinar o custo de tornar C disponível para número maior de pessoas, havendo economia de escala, se, para atender o dobro de pessoas, não for preciso o dobro de *inputs*. As tendências nos custos dos dois tipos de *outputs* podem ser muito diferentes, haja vista que as mudanças ambientais na jurisdição podem aumentar mais rapidamente os custos de fornecer C do que os custos de prover o vetor D de *outputs* diretos. Por exemplo, a presença de epidemias (E) pode tornar mais difícil a cura do paciente (C), demandando mais recursos para alcançar os objetivos de um programa de saúde, mesmo que os procedimentos de saúde se mantenham no mesmo patamar de custos.

Pela análise dos argumentos de Bradford, Malt e Oates (1969), é possível pensar em duas relações: *inputs* e *outputs* diretos e *inputs* e *outputs* finais. Um programa social pode exigir configurações diferentes em termos de ações e, consequentemente, produtos diretos, conforme as circunstâncias ambientais. Portanto, haverá gastos diferentes para alcançar os objetivos relacionados ao bem-estar da população, mesmo que o custo em si dos produtos diretos de cada ação governamental não se altere. É importante observar que a ineficiência do gasto com *outputs* diretos afeta o gasto com os *outputs* de interesse final do cidadão.

As variáveis ambientais ou exógenas poderiam também ser pensadas em termos da transformação de *inputs* em *outputs* diretos. Assim, a função de produção poderia ser representada

por $D = f(I, W)$, em que D é a quantidade máxima de *outputs* diretos obtidos, pelo vetor de *inputs-I* e influenciada pelo vetor de variáveis ambientais ou exógenas W , com base em determinada tecnologia de produção.

Uma jurisdição pode fazer mais com menos recursos ou vice-versa, dependendo de suas condições ambientais ou das variáveis exógenas. Diferenças na densidade e composição demográfica e na distribuição da população por faixa etária, por exemplo, podem afetar o valor das despesas públicas para o fornecimento de padrões comparáveis de serviços públicos. Burkhead (1971, p.187) menciona que as comparações de desempenho podem ser prejudicadas pelas condições naturais de realização do serviço; por exemplo, o número de visitas diárias por um agente de saúde na zona urbana pode não ser comparável àquele das realizadas na zona rural.

4. DATA ENVELOPMENT ANALYSIS: MODELO EM DOIS ESTÁGIOS

A teoria econômica da análise de eficiência é baseada no trabalho de Debreu (1951) e Koopmans (1951) sobre análise de produtividade. Farrell (1957) aprofundou o estudo dos dois autores com o objetivo de definir uma medida simples de eficiência da firma que poderia englobar múltiplos *inputs* (COELLI, RAO e BATTESE, 1998, p.134).

Farrell (1957, p.254) dividiu a eficiência de uma firma em dois componentes: eficiência técnica e eficiência alocativa. A primeira reflete a habilidade de uma firma obter *output* máximo dado um conjunto de *inputs*, já a segunda refere-se à capacidade de a firma usar os *inputs* em proporção ótima, considerando seus preços e a tecnologia de produção. A combinação das duas medidas fornece a medida de eficiência econômica total, a qual reflete a capacidade de uma entidade obter máximos *outputs* ao menor custo.

O cálculo das medidas de eficiência pressupõe que a função de produção seja conhecida. Na prática, isso não acontece e ela precisa ser estimada a partir de uma amostra de dados com o uso de métodos paramétricos ou não paramétricos, sendo esse último o caso da *Data Envelopment Analysis* (DEA), termo usado primeiramente por Charnes, Cooper e Rhodes (1978), constituindo uma técnica de otimização, baseada em programação linear e projetada para estabelecer uma medida de eficiência relativa entre diferentes unidades tomadoras de decisão (*Decision Making Units* – DMUs).

Os vários modelos de DEA baseiam-se na análise de eficiência das unidades tomadoras de decisão (DMUs) com múltiplos insumos (*inputs*) e produtos (*outputs*) e partem da ideia de construir uma fronteira de eficiência, em que as DMUs mais eficientes se situem sobre a superfície da fronteira, enquanto as menos eficientes se situem internamente abaixo da fronteira.

Os modelos DEA podem assumir retornos constantes de escala, ou seja, eficiência de escala, ou retornos variáveis de escala. Além disso, podem ter orientação para a maximização

de *outputs* ou minimização de *inputs*, dependendo de sob quais quantidades o administrador tem mais controle, *input* ou *output*. Como este trabalho trata do setor público de saúde e como se sabe que os burocratas ou políticos trabalham com um orçamento fixado, a orientação deve ser para os produtos, ou seja, com os insumos disponíveis, a entidade pública deve oferecer o máximo de produtos ou serviços possíveis dentro de certos padrões de qualidade.

O modelo DEA com orientação para *outputs* e retorno variáveis de escala pode ser escrito como:

$$\text{Maximize }_{\theta, \lambda} \theta \quad [1]$$

Sujeito a:

$$-\theta O_{y0} + \sum_{k=1}^n O_{yk} \lambda_k \geq 0 \quad y=1, \dots, s \quad [2]$$

$$I_{x0} - \sum_{k=1}^n I_{xk} \lambda_k \geq 0 \quad x=1, \dots, r \quad [3]$$

$$\sum_{k=1}^n \lambda_k = 1 \quad [4]$$

$$\lambda_k \geq 0 \quad [5]$$

em que:

θ = é um escalar e λ é um vetor $N \times 1$ de constantes. O valor de θ será o escore de eficiência orientado para *output* da *i*-ésima unidade de tomada de decisão;

θ = escore de eficiência relativo à DMU 0;

n = número de unidades (Total de DMUs);

s = número de *outputs*;

r = número de *inputs*;

I_{xk} = quantidade do *input* x para a DMU k ; $k=1, 2, \dots, n$;

O_{yk} = quantidade do *output* y para a DMU k ; $k=1, 2, \dots, n$.

Essencialmente, a restrição $\sum_{k=1}^n \lambda_k = 1$ assegura que uma DMU ineficiente será comparada somente a DMUs de tamanho similar. Isto é, o ponto projetado na fronteira será uma combinação convexa de DMUs observadas. Os pontos reportam-se a um segmento da fronteira ($\lambda_i + \lambda_j = 1$, em que $i \neq j$).

O escore de eficiência é dado pela medida de Farrell (1957), $1 \leq \theta < \infty$, e $\theta - 1$ é o aumento proporcional em *outputs* que se poderia alcançar pela DMU, mantendo a quantidade de *inputs* constante. Observa-se que $100/\theta$ – medida de Shephard (1970) – define o escore de eficiência que varia entre zero e 100%.

A mensuração de eficiência por meio da análise envoltória de dados gera informações úteis para o controle gerencial. Além de permitir uma análise comparativa do desempenho das DMUs a partir da captação das melhores práticas de produção,

indica a quantidade de cada *input* consumido por uma DMU ineficiente, que poderia ser reduzida ou a quantidade de cada *output* gerado que poderia ser aumentada, dependendo da orientação da análise.

No processo de transformação de *inputs* em *outputs*, como já ressaltado por Bradford, Malt e Oates (1969), existem variáveis ambientais ou exógenas que podem influenciar a eficiência de uma DMU e, portanto, deveriam ser consideradas na análise. Todavia, as variáveis exogenamente fixadas ou não controláveis, como o próprio nome sugere, não estão sob o controle dos gestores de uma DMU, no máximo eles podem administrar seus efeitos na eficiência. Consequentemente, não faz sentido falar em redução de *inputs* não controláveis ou aumento de *outputs* não controláveis.

Diante desse problema, Banker e Morey (1986) buscaram estender os modelos DEA com o intuito de estimar as reduções possíveis nos *inputs* controláveis ou discricionários pelo gestor de uma DMU, mantendo os *inputs* fixados exogenamente nos níveis correntes.

Uma alternativa para considerar as variáveis não controláveis nos modelos de avaliação de eficiência é usar o modelo DEA em dois estágios. No primeiro estágio, uma fronteira de produção é estimada a partir dos *inputs* discricionários e *outputs* de uma amostra de DMUs, sem considerar as variáveis não controláveis. No segundo estágio, os escores de eficiência são tomados como variável dependente em um modelo de regressão e as variáveis ambientais como variáveis independentes. Os sinais dos coeficientes dos termos da regressão indicam a direção da influência das variáveis não controláveis nos escores de eficiência e os testes de hipóteses padrão são usados para avaliar a força da relação. A regressão do segundo estágio também permite ajustar todos os escores de eficiência pelo uso dos coeficientes estimados em um ambiente comum.

Em um modelo orientado para *outputs*, a regressão mais simples envolve o índice de eficiência θ e as variáveis não controláveis, destacadas aqui como z_i :

$$\hat{\theta} = \beta \cdot z_i + \varepsilon_i \geq 1 \quad [6]$$

O β é um vetor de parâmetros empregados para captar a influência de z sobre o índice calculado na etapa anterior.

5. TRAJETÓRIA METODOLÓGICA

5.1. Censo de municípios paulistas

A unidade de análise deste estudo foi o município, como responsável pela subfunção atenção básica à saúde. A população, portanto, foi composta pelos municípios do estado de São Paulo, escolhidos por causa da relevância da sua economia no Brasil e pela diversidade socioeconômica, como apontado pelo Índice Paulista de Responsabilidade Social, divulgado pela Fundação Sistema Estadual de Análise de Dados – SEADE

(2008). O estudo foi realizado por meio de um censo dos municípios paulistas.

5.2. Escore de eficiência técnica

Neste primeiro estágio, uma fronteira de produção foi estimada a partir do *input* discricionário e dos *outputs* relativos aos serviços de atenção básica à saúde prestados pelos municípios paulistas.

Como condição para receber recursos do Fundo Nacional de Saúde, os governos locais são obrigados a alimentar as bases de dados administradas pelo Ministério da Saúde, dentre elas: o SIA (Sistema de Informações Ambulatoriais), o SIAB (Sistema de Informações da Atenção Básica) e o SIOPS (Sistema de Informações sobre Orçamentos Públicos em Saúde). O SIA contém informações sobre a produção ambulatorial de baixa, média e alta complexidades em termos de quantidades/valores aprovados e quantidades/valores apresentados. O SIAB possui dados gerados a partir do trabalho das Equipes de Saúde da Família (ESF) e Agentes Comunitários de Saúde (ACS). Já o SIOPS apresenta dados sobre receitas e despesas dos municípios na área de saúde e um conjunto de indicadores gerados a partir de tais informações.

Entre os 645 municípios paulistas, existem estruturas muito diferenciadas de sistemas de saúde e, conseqüentemente, dos tipos de serviços prestados. Além disso, diversas ações são de competência partilhada, envolvendo fluxo de recursos e pacientes entre municípios delimitados por uma região, macrorregião ou estado, conforme o grau de complexidade. Sendo

assim, é preciso cuidado para que não se proceda a uma análise comparativa do desempenho de estruturas heterogêneas quanto ao processo de produção em saúde.

Nesse sentido, optou-se por comparar o desempenho dos municípios paulistas quanto à produção de *outputs* diretos relativos à subfunção atenção básica à saúde. Isso porque a responsabilidade por essa subfunção é exclusiva de cada município, reduzindo as interdependências e complementaridades dos sistemas municipais de saúde e tornando possível a definição do município como um centro de *input-output*.

Para levantar e medir a eficiência relativa da aplicação de recursos na subfunção atenção básica à saúde pelos municípios paulistas, foi usada a *Data Envelopment Analysis* com orientação para a maximização de *outputs*, com o intuito de verificar o grau de reversão da função orçamento-*output*. Quanto ao rendimento de escala, análises preliminares indicaram que seria variável, o que foi seguido no restante dos procedimentos adotados.

Para os municípios paulistas, foi calculada uma função de custo composta de um insumo (*input*) e de seis produtos diretos referentes ao atendimento ambulatorial de atenção básica, ao PSF (Programa de Saúde da Família) e ao PACS (Programa de Agentes Comunitários de Saúde) (*outputs* 1 a 6), conforme pode ser visto no quadro 1.

Nessa fase, 35 municípios foram excluídos da análise por falta de dados sobre despesa com atenção básica e sete por apresentarem inconsistências nos dados sobre a produção ambulatorial em atenção básica. Ao todo, 42 observações foram eliminadas, o que representa 6,51% da população objeto de estudo.

Quadro 1

Variáveis de Input Discricionário e Outputs do Modelo de Eficiência Técnica dos Municípios Paulistas – 2006

Tipo	Variável	Descrição	Fonte
INPUT 1	Despesa com Atenção Básica	Total da despesa liquidada com a subfunção atenção básica à saúde.	SIOPS/DATASUS
OUTPUT 1	Ações de Enfermagem e Outros Procedimentos de Saúde – Nível Médio	Quantidade de procedimentos realizados neste tipo de ação.	SIA/DATASUS
OUTPUT 2	Ações Médicas Básicas	Quantidade de procedimentos realizados neste tipo de ação.	SIA/DATASUS
OUTPUT 3	Ações Básicas em Odontologia	Quantidade de procedimentos realizados neste tipo de ação.	SIA/DATASUS
OUTPUT 4	Ações Executadas por Outros Profissionais – Nível Superior	Quantidade de procedimentos realizados neste tipo de ação.	SIA/DATASUS
OUTPUT 5	Pessoas Cadastradas pelo PSF	Quantidade de pessoas cadastradas pelo Programa Saúde de Família.	SIAB/DATASUS
OUTPUT 6	Pessoas Cadastradas pelo PACS	Quantidade de pessoas cadastradas pelo Programa Agentes Comunitários de Saúde.	SIAB/DATASUS

Os escores de eficiência calculados pelo modelo DEA podem ser severamente influenciados pela presença de *outliers* nos dados, pois uma simples observação muito distante da média empurra a fronteira e, em consequência, diminui artificialmente a eficiência para o conjunto de DMUs como um todo e, assim, ocorre uma subestimação dos escores de eficiência. A base para a detecção de *outliers*, conforme proposto por Sousa e Stošić (2005), foi a medida *leverage*, termo emprestado da literatura sobre regressão linear e que indica a influência da remoção de uma dada DMU nos escores de eficiência das outras DMUs. Espera-se que os *outliers* e as DMUs com erros nos dados apresentem maior *leverage* do que o *leverage* médio global. Nos extremos, *leverage* igual a 0 (zero) significa que a remoção daquela DMU não tem efeito sobre os escores de eficiência das outras DMUs e *leverage* igual a 1 (um) indica que a remoção de determinada DMU provoca a eficiência de todas as outras, alterando os escores de zero para um.

Dez municípios foram considerados *outliers*. Analisando os dados de cada um dos municípios, individualmente, observou-se que quatro deles apresentavam um valor de despesa com atenção básica muito discrepante dos outros e, por isso, decidiu-se excluí-los da análise. Os outros seis municípios permaneceram no estudo, pois, como, aparentemente, não existiam problemas com os dados, eles poderiam servir como referência de boas práticas.

5.3. Variáveis não controláveis

Neste segundo estágio da *Data Envelopment Analysis*, buscou-se identificar variáveis que interferem e em que direções o fazem no grau de reversão da função orçamento-*output*. Existem fatores que, supostamente, influenciam a obtenção do uso eficiente dos recursos públicos que os gestores da área da

saúde, em curto prazo, não conseguem alterar, mas precisam estar atentos aos seus efeitos no desempenho de seus centros de responsabilidade.

Alguns desses fatores relacionam-se com as características populacionais. O maior percentual de população rural e/ou a menor densidade populacional podem tornar os custos da prestação de serviços em saúde mais altos por causa do maior esforço para chegar-se até a população, inclusive com perdas de economia de escala. Da mesma forma, um alto percentual da população na faixa etária acima de 60 anos pode representar altos custos de locomoção, ou do paciente até a unidade de saúde ou das equipes de saúde até a sua residência. Igualmente, vale perceber se a população na cauda inferior da faixa-etária – 0 a 18 anos – afeta os custos da prestação de serviços em atenção básica. As duas faixas etárias foram escolhidas de acordo com as tabelas utilizadas pelos planos de saúde.

Também se poderia pensar em como a organização dos estabelecimentos de saúde com maior capacidade de atendimento quanto ao volume de pessoas pode favorecer ou não o uso eficiente dos recursos públicos. Outro fator é o percentual de recursos próprios aplicados na função saúde que, por exigência legal, deve ser no mínimo de 15%. Talvez essa norma possa servir como incentivo para o gasto com desperdício, já que o governo local precisa atingir esse mínimo de qualquer maneira. Todavia, sabe-se que muitos municípios gastam além do mínimo, indicando a prioridade que dão à função saúde.

Sendo assim, fez-se a opção pelo uso da análise de regressão – $\theta_i = \beta \cdot z_i + \varepsilon_i \geq 1$ –, na qual a variável dependente foi o escore de eficiência calculado no primeiro estágio e as variáveis independentes aquelas listadas no quadro 2.

Todas as variáveis são relativas ao ano de 2006, com exceção da IND 4, cujos dados disponíveis mais recentes são de 2005.

Quadro 2

Variáveis Não Controláveis do Modelo de Eficiência Técnica dos Municípios Paulistas – 2006

Tipo	Variável	Descrição	Fonte
IND 1	População Urbana	Percentual da população urbana em relação à população total.	Fundação SEADE
IND 2	Faixa Etária I	Percentual da população com idade entre 0 e 18 anos.	DATASUS
IND 3	Faixa Etária II	Percentual da população com idade de 60 anos ou mais.	DATASUS
IND 4	Densidade Populacional	Relação entre o total da população e a área do município.	Fundação SEADE
IND 5	Escala dos Estabelecimentos de Saúde	População dividida pela quantidade de estabelecimentos ligados à gestão pública municipal que prestam serviços de atenção básica.	DATASUS
IND 6	Percentual de Recursos Próprios Aplicados na Função Saúde	Percentual de aplicação de recursos próprios aplicados em saúde calculados em conformidade com a Emenda Constitucional 29/2000.	SIOPS/DATASUS

O termo de erro da regressão do segundo estágio do modelo DEA reflete os escores de eficiência de cada município paulista ajustado pelos efeitos dos *inputs* não controláveis. Para que esse termo indique as ineficiências ajustadas, θ_i , ele deve variar de $1 \rightarrow \infty$, contemplando, assim, a noção de fronteira. De acordo com Greene (1980, p.34), com o uso de mínimos quadrados ordinários (MQO), os parâmetros são consistentemente estimados, com exceção do intercepto. Contudo, o maior erro obtido por MQO pode ser usado como estimador consistente do intercepto. Assim, para cada DMU, somando-se ao erro da regressão o erro máximo, obteve-se o índice de eficiência ajustado, pois, com o deslocamento do intercepto, todos os erros passam a ter o sinal correto. A expressão para o escore de eficiência ajustado θ_i^A é:

$$\theta_i^A = \varepsilon_i + \text{Max}_n \varepsilon_n \quad [7]$$

6. ANÁLISE DOS RESULTADOS

No primeiro estágio, a fronteira de eficiência técnica foi formada por 17 dos 599 municípios paulistas sob análise. Os outros 582 municípios encontravam-se abaixo da fronteira, ou seja, deveriam oferecer mais serviços de saúde dada a despesa com a subfunção atenção básica. Os resultados sinalizam que 2,83% dos municípios foram considerados eficientes e 97,17% ineficientes.

O escore mínimo de eficiência técnica foi de 2,62% e o máximo de 100% – medida de *Shephard*. Dos 599 municípios, 25% apresentaram escores de eficiência menores do que 13,87%, metade apresentou escores menores do que 21,27% e 25% maiores do que 32,93%, entre esses últimos estão aqueles classificados como eficientes.

Quando se analisam somente os municípios ineficientes, a média dos escores de eficiência é de 25,8%, indicando que se poderia aumentar a oferta de produtos diretos da atenção básica em 74,2% sem necessidade de direcionar mais recursos do orçamento público para a área da saúde.

Pela avaliação geral das metas de melhorias dos *outputs*, a ação dos municípios deveria estar voltada para aumentar a cobertura do PACS e do PSF e a quantidade de procedimentos ofertados por outros profissionais de nível superior e de odontologia.

No segundo estágio, os *inputs* não controláveis foram as variáveis independentes e os logaritmos naturais dos escores de *Farrell*-eficiência (o inverso da medida *Shephard*-eficiência usada no primeiro estágio), a variável dependente da análise de regressão múltipla.

A tabela 1 demonstra o resultado da análise de regressão clássica realizada com o método *stepwise*. Todos os coeficientes podem ser considerados diferentes de zero a um nível de significância de 0,05. As variáveis faixa etária I (população entre 0 e 18 anos) e percentual de recursos próprios não foram adicionadas ao modelo, indicando que, na presença das outras

Tabela 1

Resultados da Regressão do Modelo DEA em Dois Estágios

Variáveis	R ² = 0,1817		Sig.
	Coeficientes		
	Não Padronizados	Padronizados	
Constante	1,47382		0,00
Faixa Etária II	0,06528	0,2391	0,00
Densidade Populacional	-0,00010	-0,1627	0,00
População Urbana	-0,00671	-0,1381	0,00
Escala dos Estabelecimentos	-0,00001	-0,0921	0,02

variáveis, tais fatores não ajudam a explicar a eficiência da obtenção de *outputs* diretos em atenção básica.

Os pressupostos do modelo de regressão – normalidade dos resíduos e homocedasticidade – também foram testados. Os resultados apontaram para a aceitação do modelo.

O R² sinaliza que 18,17% da variação nos escores de eficiência calculados na primeira etapa podem ser explicados pelas condições ambientais dos municípios quanto à proporção da população idosa, densidade populacional, proporção da população urbana e escala de funcionamento dos estabelecimentos de saúde. Os resultados sugerem que a eficiência/ineficiência se deve, em grande parte, às decisões dos gestores públicos municipais. Entretanto, alguns deles apresentam melhor desempenho relativo ao que seria alcançado se as condições ambientais não favorecessem a prestação de serviços de saúde com menor custo. Lembrando que a medida *Farrell*-eficiência varia de 1 a ∞ , quanto maior o escore, mais ineficiente é o município. Pelos coeficientes padronizados da regressão, é possível perceber que a maior proporção de idosos na população aumenta a ineficiência de uma jurisdição e essa é a variável que mais impacta os escores de eficiência. Com o envelhecimento da população, inclusive como resultado da própria política de saúde, a tendência é que o funcionamento do sistema apresente custos mais elevados, já que o idoso apresenta restrições quanto ao deslocamento até as unidades de saúde ou, alternativamente, existem custos adicionais para que as equipes de saúde se desloquem até o paciente. Além disso, o tempo de realização dos procedimentos de saúde pode ser maior no caso da população idosa.

Por outro lado, algumas variáveis de estado ou *inputs* não controláveis favorecem a obtenção de melhores resultados pelos municípios, como acontece com a densidade populacional, o

percentual de população urbana e a escala de estabelecimentos. De fato, ter uma população menos dispersa e com maior grau de urbanização reduz o custo da prestação de serviços em atenção básica. Além disso, apesar de menos relevante, a escala com que se trabalha impacta o desempenho quanto à eficiência. Portanto, esses aspectos deveriam ser considerados nas decisões sobre a descentralização das ações de saúde.

Com base nos resultados da regressão, os escores de Farrell-eficiência foram ajustados somando-se aos termos de erro de regressão o erro máximo obtido. Para que os escores de eficiência do primeiro e do segundo estágios pudessem ser comparados, calculou-se o inverso do exponencial da medida de eficiência ajustada gerada nesse segundo estágio. O escore ajustado mínimo de eficiência técnica foi de 1,89% e o máximo de 96,45%. Dos 599 municípios, 25% apresentaram escores de eficiência ajustados menores do que 9,20%, metade apresentou escores menores do que 13,58% e 25% maiores do que 20,05%.

A tabela 2 mostra uma comparação entre os municípios eficientes do primeiro estágio, com os mais eficientes do segundo estágio.

Os municípios de Piracicaba, Rio Grande da Serra, Santo André, São Bernardo do Campo, São Caetano do Sul e São Paulo foram classificados como eficientes no primeiro estágio, mas não permaneceram no grupo dos 17 mais eficientes quando os *inputs* não controláveis foram inseridos na análise. Observa-se que, no primeiro estágio, a população dos municípios é maior do que no segundo estágio. Dos grandes municípios, permaneceram Santos e São Carlos.

Na tabela 3, mostra-se uma comparação entre os municípios mais ineficientes no primeiro e no segundo estágios.

Borá aparece como o município mais ineficiente nos dois estágios; vale ressaltar que esse é o menor município do estado de São Paulo. Com exceção de Louveira, todos os municípios possuem população abaixo de 10.000 habitantes. Nota-se que, mesmo com a exclusão do efeito de variáveis que poderiam favorecer os grandes municípios na avaliação de desempenho comparativa, somente duas jurisdições saíram do grupo dos 10 mais ineficientes: Santana da Ponte Pensa e Aspásia. Aparentemente, por suas características, tais municípios são ineficientes por natureza.

Tabela 2

Municípios mais Eficientes no Primeiro e no Segundo Estágios do Modelo DEA pela Medida de Shephard – 2006

Primeiro Estágio			Segundo Estágio		
Município	Escore de Eficiência	População	Município	Escore de Eficiência	População
Bofete	100	8.605	Tuiuti	96,45	5.593
Lins	100	71.383	Nova Guataporanga	79,69	2.053
Monte Alto	100	46.474	Sabino	74,71	5.234
Monte Mor	100	46.047	Lins	66,20	71.383
Patrocínio Paulista	100	12.672	Santos	65,47	418.375
Piracicaba	100	366.444	Monte Mor	65,42	46.047
Rancharia	100	30.147	Bofete	64,85	8.605
Ribeirão Branco	100	22.672	Rancharia	64,21	30.147
Rio Grande da Serra	100	42.406	Patrocínio Paulista	64,15	12.672
Sabino	100	5.234	Garça	63,63	44.401
Santo André	100	673.233	Monte Alto	63,06	46.474
Santos	100	418.375	Fernandópolis	63,05	65.717
São Bernardo do Campo	100	803.904	São João da Boa Vista	62,90	83.478
São Caetano do Sul	100	133.241	Ribeirão Branco	62,63	22.672
São Carlos	100	218.701	Itapira	61,59	68.398
São Paulo	100	11.016.708	São Carlos	55,28	218.701
Tuiuti	100	5.593	Vargem Grande do Sul	54,17	40.259

Tabela 3

Municípios Paulistas mais Ineficientes no Primeiro e no Segundo Estágios do Modelo DEA pela Medida de Shephard – 2006

Primeiro Estágio			Segundo Estágio		
Município	Escore de Eficiência	População	Município	Escore de Eficiência	População
Borá	2,62	826	Borá	1,89	826
União Paulista	2,87	1.376	Ocaçu	2,15	4.055
Santa Clara d'Oeste	2,89	1.846	Torrinha	2,43	9.824
Ocaçu	2,99	4.055	União Paulista	2,45	1.376
Balbinos	3,06	1.382	Balbinos	2,53	1.382
Torrinha	3,23	9.824	São João das Duas Pontes	2,59	2.560
Santana da Ponte Pensa	3,46	1.551	Paulistânia	2,77	1.965
São João das Duas Pontes	3,60	2.560	Louveira	2,80	29.552
Aspásia	3,66	1.792	Santa Clara d'Oeste	2,84	1.846
Paulistânia	3,72	1.965	Ubarana	2,84	5.240

7. CONSIDERAÇÕES FINAIS

No presente artigo, teve-se como objetivo construir uma metodologia de avaliação de desempenho dos municípios paulistas quanto à eficiência técnica na aplicação de recursos públicos nas ações de atenção básica à saúde e analisar a influência de variáveis não controláveis no processo de produção em tal área.

Para isso, foi usado o modelo DEA em dois estágios. No primeiro estágio, os escores de eficiência dos municípios paulistas foram calculados com base em um *input* discricionário – despesa liquidada com atenção básica – e seis *outputs* diretos representativos da atenção básica à saúde – número de procedimentos nesse nível de atenção e cobertura dos programas PSF e PACS. Os resultados indicaram que somente 17 dos 599 municípios sob análise foram considerados eficientes e que era possível aumentar, consideravelmente, a quantidade de serviços prestados à população sem a necessidade de aumentar as dotações orçamentárias para a área de saúde.

O maior alcance da política de saúde depende, entre outros elementos, da quantidade de *outputs* diretos (consultas, medicamentos, vacinas etc.) que se conseguem ofertar, uma vez que o consumo de tais produtos por um indivíduo inibe o consumo por outro, apesar de os benefícios indiretos poderem ser sentidos por toda a sociedade. Atualmente, o acesso ao sistema de saúde é um dos pontos críticos enfrentados pelo SUS para que se alcancem melhorias no estado de saúde da população (*output* final). Os resultados da análise comparativa de desempenho sinalizam que esse problema poderia ser, pelo menos, parcialmente resolvido com a melhor aplicação dos recursos públicos.

Como a atenção básica é uma subfunção exclusiva de cada município, é possível responsabilizar o gestor público municipal pelos resultados alcançados com a aplicação de recursos destinados a essa área. Entretanto, ao se fazer a avaliação de desempenho comparativa, um gestor não pode ser penalizado por atuar em um município com condições socioeconômicas desfavoráveis à execução dos serviços de saúde.

Sendo assim, no segundo estágio do modelo DEA buscou-se identificar as variáveis que afetam o desempenho do gestor público e a direção em que o fazem. Verificou-se que a maior proporção de idosos em uma jurisdição torna a prestação de serviços mais cara. Por outro lado, maiores densidade populacional, grau de urbanização e escala dos estabelecimentos de saúde favorecem o gasto público com eficiência. Em média, os escores de eficiência ajustados foram menores do que os originalmente calculados, indicando que as variáveis não controláveis têm um impacto líquido positivo. Isto é, baixa densidade populacional, grande parcela de população rural, maior proporção de idosos e pequena escala dos estabelecimentos de saúde são características que – apresentadas, individual ou conjuntamente, por um município – o coloca em condição desfavorável em relação aos demais. Portanto, na avaliação comparativa de desempenho, o município não deveria ser penalizado por isso.

Apesar de tais variáveis não estarem sob controle do gestor público municipal, ele não pode ignorá-las, pois afetam, diretamente, o desempenho do sistema de saúde municipal. Além do que, se for pensado que o SUS é uma política compartilhada entre os entes federativos, esses aspectos devem ser conside-

rados nas decisões conjuntas, principalmente no processo de regionalização da atenção à saúde e na pactuação de metas. Particularmente, em relação à contratualização de resultados, é desaconselhável estabelecer uma única meta para todos os municípios, já que as condições ambientais são diferentes entre eles. Por exemplo, a meta de cobertura do PSF estabelecida no pacto pela saúde pode ser facilmente alcançada por um município com alta densidade populacional, o mesmo não ocorrendo com aqueles com situação inversa.

Também é preciso pensar que a atenção básica, quando bem realizada, pode evitar diversos custos adicionais ao sistema de saúde como um todo. No sistema de saúde, a atenção básica não é vista isoladamente, mas como parte de uma rede regionalizada de assistência à saúde. As ações desenvolvidas no âmbito dessa subfunção, principalmente aquelas relacionadas à prevenção e à promoção da saúde, podem evitar a ocorrência de procedimentos de média/alta complexidade. Por exemplo, a redução de internações por diabetes *mellitus*.

Com o ajuste dos escores de eficiência, municípios considerados eficientes na primeira análise deixaram de sê-lo no segundo estágio, sinalizando que o desempenho do gestor municipal tinha sido favorecido pelas variáveis ambientais. Analisando o grupo dos mais ineficientes, percebeu-se que a composição quase não se alterou e que, predominantemente, era formado por municípios de pequeno porte. Os resultados são indicativos de que a configuração desses municípios, por natureza, leva-os à ineficiência. Portanto, cabe a indagação se municípios com população tão pequena conseguem atender a condições mínimas de escala para a oferta dos serviços de saúde e, conseqüentemente, se a atenção básica deveria ser considerada responsabilidade exclusiva de tais municípios.

Neste trabalho, focou-se a atenção básica para permitir uma avaliação comparativa de desempenho entre municípios. Entretanto, também é importante avaliar as outras subfunções de saúde, como o atendimento ambulatorial de média e alta complexidades e o hospitalar. Para isso, poder-se-iam segmentar as atividades com as mesmas características de produção entre os municípios ou segmentar os municípios conforme sua capacidade de produção. Outra forma seria avaliar as regiões ou macrorregiões de saúde como centro de responsabilidades e não o município, aparentemente a alternativa mais viável, haja vista as diretrizes do SUS quanto à organização dos serviços de saúde em redes regionalizadas.

A metodologia de avaliação de eficiência apresentada neste trabalho pode ser aplicada a outros municípios brasileiros. Além disso, cabe ressaltar que a eficiência, na área da saúde, pode ser percebida sob duas direções: recursos X *outputs* diretos e recursos X *outputs* finais. O desperdício de recursos públicos ocorre quando se incorre em custos maiores do que os necessários para realizar os procedimentos e as ações de saúde, o que se relaciona com vários casos de corrupção, por exemplo, ou, adicionalmente, quando esses procedimentos e ações não são adequadamente planejados em programas que possibilitem melhorias no estado de saúde da população. O foco deste artigo foi na primeira relação, possibilitando gerar informações sobre o que se faz com o dinheiro público.

Enfim, para além dos resultados empíricos apresentados, acredita-se que este trabalho tenha contribuído para indicar possibilidades e restrições quanto à realização de avaliações de desempenho comparativas, apresentando uma metodologia que auxilie os avanços nos estudos sobre controladoria na gestão pública. ◆

REFERÊNCIAS

- ANTHONY, R.N.; GOVINDARAJAN, V. *Sistemas de controle gerencial*. 12.ed. São Paulo: McGraw-Hill, 2008.
- BANKER, Rajiv D.; MOREY, Richard C. Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research*, Maryland, v.34, n.4, p.513-521, July/Aug. 1986. [DOI: 10.1287/opre.34.4.513].
- BORINELLI, Márcio Luiz. *Estrutura conceitual básica de controladoria: sistematização à luz da teoria e da práxis*. 2006. Tese (Doutorado em Controladoria e Contabilidade) – Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, São Paulo, Brasil, 2006.
- BRADFORD, D.F.; MALT, R.A.; OATES, W.E. The rising cost of local public services: some evidence and reflections. *National Tax Journal*, Washington, v.22, n.2, p.185-202, June 1969.
- BURKHEAD, J. *Orçamento público*. Rio de Janeiro: Fundação Getúlio Vargas, 1971.
- CHARNES, A.; COOPER, W.W.; RHODES, E. Measuring the efficiency of decision making units. *European Journal of Operational Research*, Washington, v.2, n.6, p.429-444, Nov. 1978. [DOI: 10.1016/0377-2217(78)90138-8].
- CHENHALL, Robert H. Theorizing contingencies in management control systems research. In: CHAPMAN, C.S.; HOPWOOD, A.G.; SHIELDS, M.D. (Ed.). *Handbook of management accounting research*. Oxford: Elsevier, 2007. v.1.
- COELLI, T.; RAO, D.S.P.; BATTESE, G.E. *An introduction to efficiency and productivity analysis*. Boston: Kluwer Academic, 1998. [DOI: 10.1007/978-1-4615-5493-6].
- COHEN, E.; FRANCO, R. *Avaliação de projetos sociais*. 6.ed. Petrópolis, RJ: Vozes, 2004.
- DEBREU, G. The coefficient of resource utilisation. *Econometrica*, Evanston, v.19, n.3, p.273-292, July 1951. [DOI: 10.2307/1906814].

REFERÊNCIAS

- FARRELL, M.J. The measurement of productive efficiency. *Journal of the Royal Statistical Society*, London, Series A, v.120, n.3, p.253-290, 1957. [DOI: 10.2307/2343100].
- FUNDAÇÃO SISTEMA ESTADUAL DE ANÁLISE DE DADOS (SEADE). Disponível em: <www.seade.sp.gov.br>. Acesso em: 05 mar. 2008.
- GIACOMONI, James. *Orçamento público*. 9.ed. São Paulo: Atlas, 2000.
- GREENE, W.H. Maximum likelihood estimation of econometric frontier functions. *Journal of Econometrics*, North Holland, v.13, n.1, p.27-56, May 1980.
- HOFSTEDTE, Geert. Management control of public and not-for-profit activities. *Accounting Organizations and Society*, The Netherlands, v.6, n.3, p.193-211, July 1981.
- KOOPMANS, T.C. An analysis of production as an efficient combination of activities. In: KOOPMANS, T.C. *Activity analysis of production and allocation*. Cowles Commission for Research in Economics, Monograph n.13. New York: John Wiley and Sons, 1951.
- MERCHANT, Kenneth A.; OTLEY, David T. A review of the literature on control and accountability. In: CHAPMAN, C.S.; HOPWOOD, A.G.; SHIELDS, M.D. (Ed.). *Handbook of management accounting research*. Oxford: Elsevier, 2007. v.2.
- SHAH, A. A practitioner's guide to intergovernmental fiscal transfers. In: BOADWAY, R.; SHAH, A. *Intergovernmental fiscal transfers: principles and practice*. Washington, D.C.: World Bank, 2007.
- SHEPHARD, R.W. *Theory of cost and production function*. Princeton: Princeton University Press, 1970.
- SOUSA, M.C.S. de; STOŠIĆ, B. Technical efficiency of the Brazilian municipalities: correcting nonparametric frontier measurements for outliers. *Journal of Productivity Analysis*, Norwell, v.24, n.2, p.157-181, Oct. 2005.
- VERBEETEN, Frank H.M. Performance management practices in public sector organizations: impact on performance. *Accounting, Auditing & Accountability Journal*, The Netherlands, v.21, n.3, p.427-454, 2008. [DOI: 10.1108/09513570810863996].

ABSTRACT

Performance of São Paulo State municipalities: an evaluation of the efficiency of primary healthcare

This research study aimed to build a performance assessment method to evaluate the technical efficiency of São Paulo State municipalities regarding the use of public resources for primary healthcare and to analyze the influence of non-controllable variables in the production process in this area. Technical efficiency is one of the performance assessment variables for public managers, reflecting an entity's ability to maximize outputs with the lowest consumption of inputs. Achieving this may be impaired or favored by environmental or non-controllable variables. These can give rise to biases if they are not considered in the performance assessment. Therefore, using two-stage Data Envelopment Analysis (DEA), the cities' efficiency scores were estimated and later adjusted through regression analysis. The results indicated that a considerable increase of health services for the population is possible with no added healthcare budget in most cities. Furthermore, the study found that a larger elderly population within a jurisdiction makes service delivery more expensive, while higher population density, urbanization level and scale of health establishments enhance efficient public spending. The five cities that were found to be most efficient were: Tuiuti, Nova Guataporanga, Sabino, Lins and Santos.

Keywords: controllership, performance evaluation, efficiency, primary health care, municipalities.

RESUMEN

Desempeño de los municipios del estado de São Paulo: una evaluación de la eficiencia de la atención básica a la salud

El objeto en este trabajo fue el de construir una metodología de evaluación de desempeño de los municipios del estado de São Paulo en lo que se refiere a la eficiencia técnica en la asignación de recursos públicos en las acciones de atención básica a la salud, y analizar la influencia de variables no controlables en el proceso de producción en dicha área. La eficiencia técnica es uno de los parámetros de evaluación de desempeño de los gestores públicos y refleja la capacidad que tiene una entidad de obtener los máximos resultados con el menor consumo de inputs. El alcance de tal métrica puede verse perjudicado o favorecido por las variables ambientales o no controlables que, si no se consideran en la evaluación de desempeño, pueden generar sesgos. En este sentido, por medio de la metodología *Data Envelopment Analysis* (DEA) en dos fases, se estimaron y luego se ajustaron las clasificaciones de eficiencia de los municipios con el uso del análisis de regresión. Los resultados indicaron que sería posible aumentar considerablemente la cantidad de servicios prestados a la población sin necesidad de nuevas partidas presupuestarias en la mayoría de los municipios. Además, se verificó que la mayor proporción de ancianos en una jurisdicción hace que la prestación de servicios sea más cara; por otro lado, mayor densidad poblacional, grado de urbanización y escala de los establecimientos de salud favorecen el gasto público eficiente. Los cinco municipios considerados más eficientes fueron Tuiuti, Nova Guataporanga, Sabino, Lins y Santos.

Palabras clave: control de gestión, evaluación de desempeño, eficiencia, atención básica a la salud, municipios.