

Rankings em Educação: Tipos, Problemas, Informações e Mudanças: Análise dos Principais Rankings Oficiais Brasileiros[†]

Eduardo de Carvalho Andrade^{*}

Resumo

Este trabalho classifica os diferentes tipos existentes de rankings na educação, bem como aponta problemas encontrados e apresenta contribuições para resolução dos mesmos. Sobre os principais rankings oficiais brasileiros conclui-se que: (i) O MEC tem uma política distinta para o ranking no Ensino Superior vis-à-vis ao do Ensino Médio – no primeiro ele estima o valor adicionado e preocupa-se em inibir comportamentos oportunistas; (ii) foco nos resultados do ENEM/ENADE pode levar a um alvo diferente da qualidade da educação; (iii) a escolha *ad hoc* dos pesos componentes do ranking da graduação torna questionável a sua utilização; (iv) é importante o fornecimento dos intervalos de confiança dos valores adicionados estimados.

Palavras-Chave

rankings, valor adicionado, comportamento oportunístico

Classificação JEL

A20, I20

Rankings in Education: Types, Problems, Information and Changes Analysis of the Most Important Brazilian Official Rankings

Abstract

This paper classifies the different types of rankings in education, shows the problems that it has been finding and presents contributions to solve them. About the most important Brazilian official rankings it concludes that: (i) the Brazilian Ministry of Education has a different policy with respect to the higher education ranking vis-à-vis the secondary one – in the latter, it estimates the value added and inhibits opportunistic behavior; (ii) the focus in the ENEM/ENADE results may not imply greater education quality; (iii) the *ad hoc* choice of the higher education ranking components weights questions its validity; (iv) it is important to provide the confidence intervals of the estimated value added.

Keywords

rankings, value added, opportunistic behavior

JEL Classification

A20, I20

[†] Artigo recebido em dezembro de 2009 e aceito para publicação em junho de 2010. O autor gostaria de agradecer os comentários do parecerista anônimo, bem como o excelente trabalho de assistência de pesquisa de Bruno Cardoso. Os erros remanescentes são de inteira responsabilidade do autor.

^{*} Professor e Pesquisador do Insper - Endereço para contato: Rua Quatá 300 - São Paulo CEP: 04546-001 - Email: eduardo.andrade@insper.edu.br

1 Introdução

Tendo avançado na quase universalização do Ensino Fundamental, a prioridade da política educacional tem se voltado para a melhoria na qualidade do ensino. Esse foco é particularmente importante quando se tem em mente que a qualidade - e não a quantidade da educação - é a variável mais importante para explicar crescimento econômico. Hanushek e Woessmann (2008) comparam os efeitos das duas variáveis, qualidade e quantidade da educação, sobre a taxa de crescimento do PIB *per capita*. O resultado indica que um aumento de um desvio-padrão da quantidade da educação, mensurado pela média dos anos de escolaridade da população, eleva o PIB *per capita* em 0,26% ao ano. Em contraste, o mesmo aumento de um desvio-padrão da qualidade da educação, medido pelo resultado no teste de proficiência em matemática, pode incrementar o PIB *per capita* em 1,4% ao ano.

Esse foco na qualidade tem levado o governo federal, através do Ministério da Educação (MEC), a tentar mensurar a qualidade da educação tanto no Ensino Básico como no Ensino Superior. Com este fim, ele tem desenvolvido métricas e divulgado rankings das instituições de ensino. O objetivo é fornecer informações sobre a qualidade das instituições para os diferentes *stakeholders*. Elas podem auxiliar pais/estudantes na escolha de onde estudar, empregadores na contratação de novos funcionários, governos nos seus programas voltados para a melhoria da qualidade da educação e o próprio MEC no seu trabalho de regulação do setor educacional.

Em relação aos cursos de graduação, o MEC instituiu, a partir de meados da década de 90, um sistema de avaliação nacional dos cursos de graduação brasileiros. Depois de sofrer várias alterações na sua forma de avaliação, atualmente, cada curso recebe uma nota entre 1 e 5. Esta nota depende de uma combinação de três fatores: (1) a média das notas dos estudantes no Exame Nacional de Desempenho dos Estudantes (ENADE); (2) os insumos utilizados no processo educacional – qualidade da infraestrutura e instalações físicas, corpo docente e recursos didático-pedagógicos; e (3) quanto o curso agrega ao aprendizado do aluno, ou o valor adicionado, tecnicamente chamado de Indicador de Diferença de Desempenho.¹

Além deste ranking oficial, existe um ranking privado, que é o mais tradicional dos cursos da graduação e é publicado anualmente pela revista *Guia do Estudante*. Cada curso recebe uma avaliação, podendo ser cinco, quatro, três ou nenhuma estrela. O ranking é construído a partir da opinião de pareceristas contratados pela revista.² Esses pareceristas são professores que recebem informações sobre o corpo docente e as publicações acadêmicas de cada instituição.

1 Para maiores detalhes sobre a metodologia, ver a nota técnica do INEP (2008).

2 Para maiores detalhes sobre a metodologia, ver *Guia do Estudante* (2008).

No que se refere ao Ensino Médio, existe somente um ranking das escolas: ele é construído a partir dos resultados dos seus alunos no Exame Nacional do Ensino Médio (ENEM). Esta é uma prova oficial, não obrigatória, cujo resultado é utilizado por muitas IES no seu processo de seleção dos novos alunos. Cada escola recebe uma nota, que é igual à média das notas de todos os seus alunos que prestam o exame. As escolas com as maiores médias são as melhores colocadas no ranking.

Além de uma forte repercussão na mídia, a divulgação dos rankings tem impactos e consequências importantes. Andrade *et al.* (2009) mostram que a posição de uma Instituição de Ensino Superior (IES) no *Guia de Estudante* afeta de forma significativa o seu *market share*. Em particular, a mudança no ranking passando de quatro para cinco estrelas permite a instituição aumentar a sua mensalidade em R\$ 486 reais sem perder a sua fatia de mercado. No mesmo trabalho, os autores mostram que a nota do ENADE dos ingressantes, que é um dos componentes do ranking oficial, também afeta significativamente o *market share* de uma IES. O aumento de um ponto na nota do ENADE, que tem uma escala de 0 a 100, permite à IES aumentar sua mensalidade em R\$ 16 reais, sem perder *market share*.³

Talvez como reflexo destes impactos, é comum ver nos folhetos de divulgação, particularmente das instituições privadas, as posições dos cursos nos rankings. Mais recentemente, as mesmas iniciaram um movimento junto ao Congresso Nacional para que o MEC seja proibido de divulgar os dados das suas avaliações, possivelmente devido ao impacto financeiro dos maus resultados.⁴

Não existem estudos semelhantes sobre o impacto do ranking das escolas de Ensino Médio sobre as mesmas. No entanto, a evidência anedótica sugere que os seus impactos podem ser também importantes. Os dirigentes das escolas mal colocadas no ranking são questionados pelos pais dos alunos, que exigem explicações e estratégias para melhorar a posição das escolas. A procura pelas escolas melhores ranqueadas parece também ser maior.

Dado os impactos e a importância dos rankings, é importante examinar com detalhes os diferentes tipos existentes, suas características e os problemas de cada um deles. Uma análise desta natureza tem um papel importante de orientar os diferen-

3 Impactos importantes também ocorrem nos Estados Unidos. Monks e Ehrenberg (1999) utilizam dados de painel para avaliarem o impacto do ranking da revista U.S News and World Report, o mais tradicional do mercado americano, sobre as universidades. Eles concluem que uma pior posição no ranking é prejudicial para a universidade: uma menor parcela dos alunos aceitos pela universidade decide se matricular; a qualidade das novas turmas, mensurada pela média no teste exigido para admissão, cai; o preço líquido pago pelo aluno é menor porque a universidade tem de ser mais generosa na concessão de ajuda financeira para atrair alunos dentro do menor grupo de candidatos que deseja lá estudar.

4 Ver artigo "Particular quer Fim da Divulgação de Avaliações", no jornal *O Estado de São Paulo* de 5/12/2008.

tes públicos (pais/estudantes, empresas, governos, instituições de ensino) interessados em entender como interpretar e processar as informações contidas nos diferentes rankings, bem como identificar as suas limitações. Além desses objetivos, este trabalho propõe mudanças de forma a aperfeiçoar os rankings oficiais existentes.⁵

Este trabalho divide-se em cinco seções, incluindo esta introdução. Na segunda seção serão apresentados os principais tipos de rankings de instituições de ensino. Em seguida, os principais problemas de cada tipo de ranking são examinados. Com base na análise realizada nas seções anteriores, a quarta seção examina os dois principais rankings oficiais existentes: um para as escolas de Ensino Médio e outro para as IES. O objetivo é chamar a atenção para os seus principais problemas, apontar as suas principais contribuições e sugerir mudanças com o intuito de melhorá-los. A última seção está reservada para considerações finais.

2 Tipos de Ranking

A educação pode ser entendida como um serviço que transforma insumos em produtos. Ela pode ser representada por uma função de produção, que mostra a relação existente entre esses insumos e produtos e, adicionalmente, corresponde também ao máximo de produção que pode ser alcançado dado os insumos utilizados. Em uma formulação bem geral, uma função de produção ("f") pode ser caracterizada pelo modelo descrito na seguinte equação:

$$Y_{it} = f(B_i^{(t)}, P_i^{(t)}, S_i^{(t)}, I_i), \quad (1)$$

onde: (i) Y_{it} é o produto⁶ na sua versão escalar (ou uma matriz na versão com múltiplos produtos) relacionado ao estudante "i" no tempo "t"; (ii) $B_i^{(t)}$ é um vetor das características do background familiar do estudante "i" acumulado no tempo "t"; (iii) $P_i^{(t)}$ é o vetor de influências dos pares acumulados no tempo "t"; (iv) $S_i^{(t)}$ é o vetor de insumos controlados diretamente pela escola do estudante "i" acumulado no tempo "t"; e (v) I_i é o vetor de habilidades inatas do estudante "i".⁷ Normalmente, assume-se que mais insumos acarretam maior produção, mas este aumento ocorre a taxas cada vez menores.

5 Este trabalho não analisa rankings de cursos de Pós-Graduação, somente instituições do Ensino Médio e de Graduação.

6 O produto, por exemplo, pode ser o resultado do aluno num teste de proficiência, que mensura a sua habilidade cognitiva.

7 Essa versão da função de produção é extraída de Hanushek (1979) e sinaliza que valores passados dos insumos, através do acumulado no tempo, afetam a produção, possivelmente com impactos menores com o passar do tempo.

Uma versão alternativa deste modelo é comumente utilizada. Ao invés de usar o valor do produto num determinado ponto no tempo, o foco da análise é no valor adicionado. Ou seja, quanto que ocorreu de ganho no produto de um período para outro. Assumindo que a equação (1) acima vale para diferentes pontos no tempo, o tempo “t” e um anterior “t*”, é possível levar em consideração a mudança do nível do produto entre os períodos e escrever a seguinte especificação:

$$Y_{it} - Y_{it^*} = f^*(B_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)}, I_i). \quad (2)$$

Tendo como base esse referencial teórico da função de produção, é possível fazer uma classificação dos diferentes tipos de rankings de instituições de ensino. Eles podem ser classificados em quatro tipos: foco no produto e objetivo, foco no produto e subjetivo, foco no insumo e objetivo e foco no insumo e subjetivo.⁸ Vejamos cada um deles separadamente.

2.1 Foco no Produto e Objetivo

O ranking com foco no produto e objetivo é aquele que escolhe, dentre aqueles possíveis produzidos pela instituição de ensino, algum produto (ou uma composição de produtos) que possa ser mensurado de forma objetiva. As instituições são ranqueadas em função do valor da sua produção, sendo consideradas as melhores aquelas com maior produção.

O produto pode ser entendido como o valor absoluto ou valor adicionado, respectivamente, como nas equações (1) e (2) acima. É bastante comum que o produto utilizado seja o resultado médio dos alunos da instituição de ensino num teste de proficiência. Mas várias outras alternativas poderiam ser utilizadas. Por exemplo, no caso de IES, o valor dos salários ou o grau de empregabilidade dos seus ex-alunos seriam algumas alternativas possíveis. No caso de instituições de Ensino Médio, algumas opções seriam a taxa de aprovação nos vestibulares de acesso às IES ou o desempenho posterior no curso superior.

Existem dois deste tipo de ranking no Brasil. Como mencionado anteriormente, o primeiro é o ranking das escolas do Ensino Médio, divulgado pelo MEC. Cada escola recebe uma nota que corresponde à nota média que os seus alunos obtêm no ENEM e são classificadas conforme esta nota. Neste caso, o produto é identificado pela habilidade cognitiva adquirida pelo aluno, refletida na sua nota na prova.

⁸ Esses são os rankings puros. Existem rankings que são combinações desses tipos, como o do MEC para as IES.

Espera-se que melhores notas reflitam, pelo menos, uma maior capacidade dos alunos enfrentarem suas atividades e nestas serem bem sucedidas após a conclusão do curso.

O segundo ranking é utilizado também pelo MEC para comparar as IES. Dois dos três componentes do ranking do MEC são baseados nas notas dos alunos no ENADE. Esses componentes correspondem a 70% do ranking total do MEC. No primeiro componente (com peso de 40% no ranking), é computada a média das notas dos ingressantes e concluintes no ENADE. No segundo componente (peso de 30%), é estimado o valor adicionado da instituição, em outras palavras, o cálculo da efetiva contribuição da instituição de ensino para o aprendizado do aluno.⁹ Neste caso, o aprendizado é entendido como um melhor desempenho no exame. Em ambos os componentes do ranking do MEC, é utilizado um indicador do produto da educação, ambos associados à nota no ENADE, que é um indicador objetivo.

2.2 Foco no Produto e Subjetivo

O segundo tipo de ranking é aquele com foco no produto e subjetivo. Ele caracteriza-se pela escolha, dentre aqueles possíveis produzidos pela instituição de ensino, de algum produto (ou uma composição de produtos) que possa ser mensurado. No entanto, diferente do tipo de ranking anterior, a métrica utilizada não é objetiva e sim subjetiva. As informações a respeito do produto (ou produtos) são obtidas através de entrevistas junto a diferentes públicos: alunos, ex-alunos, administradores de instituições de ensino, pareceristas contratados pelo organizador do ranking ou professores. Implicitamente, assume-se que existe uma relação forte entre as opiniões fornecidas pelos diferentes públicos e o verdadeiro valor do produto educacional. Sem dúvida, a grande vantagem dessa abordagem é que dados e informações podem ser obtidos de forma razoavelmente econômica em larga escala.

O ranking brasileiro mais tradicional dos diferentes cursos de graduação das IES encaixa-se perfeitamente neste tipo baseado no “produto e subjetivo”, qual seja, o divulgado pela revista *Guia do Estudante*. Neste ranking, todo curso que possua uma turma formada há pelo menos um ano participa da avaliação. Pareceristas, professores, coordenadores de curso e avaliadores do MEC são contratados pela revista e recebem informações sobre cada curso a ser avaliado, tais como titulação do corpo docente e produção científica. Em seguida, eles atribuem uma nota para

9 Segundo o INEP (2008), para o cálculo do valor adicionado, “considera-se o desempenho médio obtido no ENADE pelos concluintes de cada curso e o desempenho médio que seria esperado ao final do curso para o perfil de ingressantes daquela instituição (...)” (pág. 3). A diferença entre o efetivo e o esperado é a medida do valor adicionado.

o curso avaliado: excelente (cinco estrelas), muito bom (quatro estrelas), bom (três estrelas), regular, ruim ou “prefiro não opinar”. O conceito final que cada curso recebe (cinco, quatro, três ou nenhuma estrela) é em função da média das notas dadas pelos pareceristas. Esse número final captura a visão dos avaliadores sobre a qualidade do curso oferecido, no produto (ou composição de produtos) considerado por eles o mais relevante.^{10,11}

2.3 Foco no Insumo e Objetivo

O terceiro tipo de ranking é aquele com foco no insumo e objetivo. Ao contrário dos dois rankings anteriores, nesse tipo não se escolhe uma medida de produção para classificar as instituições de ensino. O foco aqui é nos insumos utilizados no processo produtivo da educação, que podem ser mensurados objetivamente.

Nessa abordagem, implicitamente, assume-se que os insumos são bons sinalizadores da produção educacional ou da qualidade de uma instituição de ensino. Em outras palavras, quanto maiores ou melhores os insumos utilizados, maior e/ou melhor vai ser a produção.¹² Por exemplo, espera-se que maiores gastos por aluno por parte da instituição, corpos docente e discente de melhor qualidade e uma melhor infraestrutura, contribuam para que os alunos formados sejam mais produtivos no mercado de trabalho, tenham melhores desempenhos nos testes de proficiência ou uma maior consciência social.

O uso deste tipo de abordagem é bastante comum. Talvez isto se explique por dois motivos: primeiro, pela maior facilidade em conseguir informações sobre os insumos utilizados, do que o produto produzido por uma instituição de ensino; segundo, provavelmente devido à existência de um menor consenso entre os especialistas sobre qual deve ser a variável (ou variáveis) utilizada como produto de um serviço educacional. Utiliza-se então, como alternativa, variáveis de insumo. Um dos componentes do atual ranking do MEC das IES é composto exatamente por informações sobre os insumos educacionais mensurados de forma objetiva. 30% do ranking são construídos a partir de informações sobre os insumos. Destes 30%, 62,7% correspondem a informações objetivas dos insumos:¹³ percentual de profes-

10 Para maiores detalhes sobre a metodologia, ver *Guia do Estudante* (2008).

11 O ranking mais tradicional dos cursos de graduação dos Estados Unidos, feito pela revista *US News and World Report*, também tem um importante componente baseado no “produto e subjetivo. Ver Clarke (2002).

12 Formalmente, também na equação (2) como na (1), assume-se que a primeira deriva da função de produção em relação a cada um dos insumos é positiva. Na próxima seção, discutem-se algumas evidências empíricas se efetivamente mais insumos significa mais e melhor produção.

13 O restante (37,3%) corresponde a informações subjetivas sobre os insumos. Ver discussão abaixo.

sores doutores no curso (38,9%) e percentual de professores que cumprem regime parcial ou integral (não horista) no curso (23,8%).¹⁴

2.4 *Foco no Insumo e Subjetivo*

O ranking com foco no insumo e subjetivo diferencia-se do anterior única e exclusivamente pelo fato das informações a respeito dos insumos serem subjetivas e não objetivas. Elas são obtidas através de questionários respondidos através de entrevistas feitas junto a diferentes públicos, por exemplo, alunos ou ex-alunos.

Obviamente a hipótese implícita, assim como no ranking anterior, é que mais insumos incrementam a qualidade ou quantidade da produção. Adicionalmente, assume-se que haja uma relação forte entre as informações fornecidas pelos provedores das informações e os verdadeiros valores dos insumos.

Esse tipo de ranking é certamente menos utilizado. Uma explicação provável é a seguinte: já que se vão obter informações subjetivas sobre um curso junto com diversos públicos, com vistas a construir um ranking, parece mais razoável obter informações diretamente sobre o produto e não sobre o insumo, que só indiretamente está relacionado com o produto, o objetivo final.

Um dos componentes do ranking do MEC das IES é composto pela qualidade dos insumos, que corresponde a 30% do total. Desse percentual, 37,3% está relacionado com duas variáveis de insumo obtidas de forma subjetiva, através de questionário respondido pelos próprios alunos. As variáveis referem-se à qualidade dos recursos didático-pedagógicos utilizados e da infraestrutura disponível.

3 Problemas dos Rankings

Nesta seção serão discutidos os problemas dos diferentes tipos de ranking.

3.1 *Foco no Insumo*

Os rankings construídos a partir das variáveis de insumo partem da hipótese implícita de que a qualidade educacional de uma instituição está necessariamente associada com mais ou melhores insumos utilizados no processo produtivo. Esta visão

¹⁴ O ranking da *US News and World Report* incorpora uma gama variada de insumos objetivos no seu ranking. Ver Clarke (2002).

está em linha com o arcabouço teórico explicitado na seção anterior da existência de uma função de produção em educação.

O principal problema deste tipo de ranking é que a hipótese implícita pode não ser correta. Por um lado, do ponto de vista teórico, não estão claros quais são os insumos, diretamente sob o controle das instituições de ensino, que devem ser utilizados como *proxy* do produto (ou produtos) educacional. Por outro lado, do ponto de vista empírico, não existe evidência de que aqueles insumos normalmente considerados relevantes para explicar o produto educacional são de fato relevantes, pelo menos quando se utiliza o resultado em testes de proficiência como medida de qualidade educacional.

No nível dos Ensinos Fundamental e Médio, Hanushek (2003) faz uma extensa revisão da literatura, para os Estados Unidos e internacional, concluindo que os insumos não explicam o desempenho escolar dos alunos, medido pelos resultados nos testes de proficiência. Para o caso brasileiro, Menezes-Filho (2007) mostra que os insumos diretamente sob o controle da instituição de ensino, tais como número de computadores na escola, escolaridade, idade e salário dos professores, têm efeitos muito reduzidos sobre o desempenho dos alunos.

No nível de graduação, o INEP (2008) mostra que diversas variáveis de insumo, como a opinião dos alunos sobre a qualidade da biblioteca, da infraestrutura geral ou a disponibilidade de computadores, não são capazes de explicar o valor adicionado da instituição. Outras, no entanto, como o percentual de professores doutores e não horistas, têm impacto significativo.

Em suma, com base no exposto acima, é questionável a adequabilidade da utilização de variáveis de insumo para comparar a qualidade das diferentes instituições de ensino. A questão central é a dificuldade de saber exatamente *a priori* quais são os insumos de fato relevantes para explicar o produto educacional. Aqueles insumos comumente considerados relevantes podem de fato não serem.

3.2 *Subjetivo*

No que diz respeito aos rankings subjetivos, as críticas podem ser divididas entre aquelas informações fornecidas por *experts* (professores, administradores e pareceristas) e por alunos.

No caso dos *experts*, a hipótese implícita de construir um ranking com base nas informações por eles fornecidas é que eles são de fato capazes de avaliar adequadamente a qualidade de todas as instituições de ensino participantes. Essa hipótese

é questionável. É pouco provável que alguém conheça e seja capaz de avaliar adequadamente os diferentes aspectos das instituições de ensino de diferentes perfis, características focadas em diferentes tipos de alunos.

Uma tentativa de resolver esse problema é fazer com que as instituições sejam avaliadas por diferentes indivíduos. Neste caso, as instituições podem, em tese, ser avaliadas por indivíduos que conhecem as suas efetivas características e qualidade. Essa potencial solução, na verdade, cria outro problema: compatibilizar num único ranking as respostas dadas por diferentes indivíduos sobre as diferentes instituições. Isto porque eles não necessariamente têm o mesmo grau de exigência, nem avaliam da mesma maneira até mesmo uma informação objetiva em comum e nem atribuem o mesmo valor para as distintas características das instituições.

Por exemplo, a revista *Guia do Estudante*, que contrata pareceristas para avaliarem as instituições, reconhece este problema. Na seção metodológica, ela escreve:

A distribuição dos cursos aos pareceristas de cada área é feita eletronicamente e de forma aleatória, e a cada consultor recebe os cursos da mesma região onde leciona. 'Isso pode até criar um viés regional caso os consultores de uma região sejam sistematicamente mais positivos – ou negativos – em suas avaliações do que a média geral. Mas, como em qualquer pesquisa de opinião, devemos levar em conta a cultura e os padrões de cada localidade'.¹⁵

Outro problema de utilizar informações subjetivas de professores, administradores e avaliadores em geral, é mencionado por Chun (2002). Segundo ele, "(...) it is generally assumed that reputations change more slowly than real changes in institutions, thus overvaluing institutions that, in fact, may be declining and undervaluing institutions that are improving".¹⁶ Ou seja, melhoras internas podem estar sendo implementadas numa determinada instituição, com impactos de curto prazo sobre a qualidade do serviço educacional oferecido, mas só posteriormente afetando a sua reputação e a avaliação dos avaliadores. Pelo menos durante um período de tempo, tal instituição seria subavaliada e sendo prejudicada por isto. Obviamente que o contrário pode ocorrer.

Existem problemas também quando o grupo responsável por fornecer as informações subjetivas são alunos. Neste caso, os estudantes oferecem informações sobre a sua instituição de ensino e as respostas são comparadas entre as diferentes instituições.

¹⁵ Trecho extraído da página 15 do *Guia do Estudante* (2008).

¹⁶ Trecho extraído da página 19 de Chun (2002).

Um dos problemas desta abordagem é semelhante àquele mencionado acima para o caso das informações serem dadas por professores, pareceristas ou administradores, ou seja, o viés das informações. Os alunos se autosselecionam entre as diferentes instituições. Isto faz com que cada instituição tenha alunos com características específicas, que são bem distintas daquelas dos estudantes das demais instituições. No mínimo os seus valores, a visão crítica ou as suas prioridades são diferentes. Isto acarreta um viés nas respostas com impactos relevantes sobre a posição de uma escola num ranking. Por exemplo, uma instituição que atraia estudantes relativamente mais críticos seria certamente prejudicada numa comparação com outras instituições, quando as informações subjetivas dos alunos definem a avaliação.

Adicionalmente, Chun (2002) fornece outro argumento que leva ao questionamento sobre a validade de rankings baseados nas respostas dos estudantes: eles podem não ter interesse em responder os questionários com informações precisas. Por exemplo, ele menciona sobre o *halo effect*, no qual “students may inflate reporting of their behavior, performance, or what they perceive they have gained from their college experience towards the more socially acceptable. (...) because the effect is consistent across students and institutions, comparisons are not compromised. (This, however, is still a concern when it comes to having an ‘accurate’ picture of student growth)”. É questionável se os resultados não ficam de fato comprometidos. Isto porque, como os alunos se autosselecionam entre as diferentes instituições, é possível que o viés dos diferentes grupos de alunos não tenha a mesma força ou intensidade.

Em suma, os rankings baseados em informações subjetivas apresentam problemas de difícil solução. Por um lado, incapacidade de um avaliador ter conhecimento sobre todas as instituições classificadas; por outro lado, viés de diferentes naturezas na classificação das instituições. Por fim, defasagem nas informações disponíveis dos avaliadores.

3.3 *Produto*

A grande dificuldade de um ranking baseado no produto (ou produtos) é a escolha do indicador apropriado. Como discutido anteriormente, a produção educacional é multifacetada e vários aspectos podem ser levados em consideração, tais como a performance em testes de proficiência, o desempenho futuro no mercado de trabalho, a consciência política, os valores morais ou o valor intrínseco do conhecimento.

A avaliação das instituições de ensino passa a ser parcial quando se escolhe um indicador ou uma composição deles, o que é inevitável de ser feito. Como consequência,

a ordem de um ranking pode variar em função do produto escolhido. Isto porque uma instituição pode naturalmente ser melhor neste indicador (em detrimento de outros) ou focar as suas ações para alcançar um melhor desempenho neste indicador específico de forma a ser melhor classificada no ranking. O fato é que não existe e provavelmente nunca existirá um consenso sobre qual produto utilizar.

Nos Ensinos Fundamental e Médio a abordagem tradicional na literatura é identificar a produção educacional como o resultado nos testes de proficiência, normalmente em matemática e língua. Sobre esse assunto, Cunha *et al.* (2005) escrevem:

Current systems of evaluating educational reforms are based predominantly on changes in scores on cognitive tests. These tests capture only one of the many skills required for a successful life. A more comprehensive evaluation of educational systems would account for their effects on producing the noncognitive traits that are also valued in the market. Muito provavelmente pela dificuldade de mensuração das habilidades não cognitivas, essas variáveis não são ainda utilizadas.

No entanto, mesmo sendo uma avaliação parcial, pode ser interessante utilizar essa métrica para ranquear as escolas, dado o impacto significativo para a sociedade e para os estudantes de um melhor desempenho nos testes de proficiência. Por exemplo, Hanushek e Woessmann (2008) mostram que existe uma forte evidência empírica de que a habilidade cognitiva de uma população (mensurada pelos resultados nos testes de proficiência) está fortemente relacionada com os ganhos dos indivíduos, a distribuição de renda e o crescimento econômico do país.¹⁷

No caso do Ensino Superior, dada a sua natureza mais complexa, a dificuldade é maior para se encontrar uma métrica para ranquear as instituições. Segundo Hanushek (1979), *few people believe that test scores adequately measure outputs* no Ensino Superior. Ele acrescenta: *Consideration of multiple outputs does suggest that production functions estimated with test-score measures might be more appropriate in earlier grades, where the emphasis tends to be more on basic cognitive skills – reading and arithmetic – than in later grades.* Portanto, rankings das IES que utilizam resultados em testes de proficiência devem ser olhados com um maior ceticismo.

Mesmo que se aceite o resultado do teste de proficiência como indicador da produção educacional, pelo menos nas escolas, isto não encerra as dificuldades com esse tipo de ranking.

¹⁷ Para mais evidências sobre os impactos dos resultados nos testes de proficiência para a sociedade e os indivíduos, ver Hanushek e Kimko (2000), Murphy e Peltzman (2004) e Bishop (1989).

A primeira dificuldade está relacionada com a questão de qual função de produção utilizar, a do valor absoluto (Equação (1)) ou a do valor adicionado (Equação (2)). Isto porque, como discutido acima, vários fatores, fora do controle das escolas, afetam o desempenho dos alunos: *background* familiar, habilidade do indivíduo, fatores históricos da formação educacional do aluno e a qualidade/formação dos colegas de classe. Portanto, é perfeitamente possível que a contribuição ou valor adicionado de uma escola para o aprendizado seja maior do que a de outra escola, mesmo sendo a média dos seus alunos inferior.

A questão então seria qual tipo de ranking é o mais apropriado. A resposta depende de qual é o objetivo esperado, pois ambos podem conter informações relevantes. O ranking baseado no valor absoluto sinaliza que talvez a escola tenha sido mais eficaz em selecionar os melhores alunos. Caso o *peer effect* seja relevante, pais/estudantes podem preferir basear neste ranking a sua tomada de decisão em qual escola estudar. Adicionalmente, para os empregadores interessados em contratar os melhores alunos, esse também é o ranking relevante. O importante para eles é contratar o profissional com o maior potencial, independentemente de como este adquiriu a sua competência.

No entanto, um ranking baseado no valor absoluto pode refletir um comportamento oportunístico por parte das instituições. A estratégia para ser bem ranqueada pode ser a seguinte: como a classificação é em função da média dos seus alunos nos testes de proficiência, pode ser de interesse da instituição evitar que os seus piores alunos realizem a prova. Sem estes, a nota da escola aumenta e, conseqüentemente, a sua posição no ranking. A maneira de conseguir isto é expulsar ou induzir os piores alunos a abandonar a escola. Alternativamente, o processo de seleção pode ser bastante rigoroso, de maneira a não aceitar alunos com baixo potencial de aprendizado e os seus impactos negativos sobre a nota da escola. Portanto, ranking baseado no valor absoluto pode induzir as escolas a adotarem estratégias não inclusivas, sem a preocupação de recuperação dos alunos com maiores dificuldades de aprendizado. Neste caso, escolas que não adotam este tipo de política podem ser prejudicadas por obterem posições inferiores no ranking, mesmo que de fato sejam mais eficientes em produzir o serviço educacional.

O ranking baseado no valor adicionado pode corrigir essa distorção e ser o mais adequado, principalmente se o objetivo é comparar a contribuição efetiva da escola. Nesse sentido, para os órgãos responsáveis pela fiscalização do serviço educacional oferecido, uma classificação das instituições de ensino baseada no valor adicionado é o mais relevante. Por exemplo, numa política educacional na qual professores, diretores e funcionários recebam bonificações ou sejam penalizados em função da sua contribuição para o produto educacional, o uso da métrica do valor adicionado

parece o mais adequado.¹⁸ Por fim, pais/alunos podem preferir uma escola que contribua mais significativamente para o aprendizado do aluno, com o maior valor adicionado possível. O certo é que instituições malranqueadas em função do seu valor adicionado sinalizam que algo precisa ser feito para tornar o seu trabalho mais efetivo em aumentar o aprendizado dos alunos.

Apesar da utilidade de um ranking baseado no valor adicionado, a qualidade da informação por ele fornecida pode ser questionada por pelo menos dois motivos: o primeiro deles está relacionado com a dificuldade de se comparar o valor adicionado de escolas com diferentes perfis de alunos. Um exemplo facilita a clarificação desse ponto. Imagine duas escolas iguais, “A” e “B”, com a mesma eficiência na produção do serviço educacional, mas se diferenciando apenas pelo fato da primeira ter um corpo discente mais qualificado e, como resultado, maiores médias nos testes de proficiência. Caso seja mais difícil aumentar o nível de aprendizado dos melhores alunos, por já partirem de um patamar superior, a escola “A” estaria abaixo da escola “B” num ranking baseado no valor adicionado. A colocação se reverteria caso seja mais difícil elevar o aprendizado dos alunos relativamente piores, que possuam um nível inicial menor. Esse é um exemplo no qual a classificação seria independente da eficiência e do trabalho realizado pelas escolas.

Já o segundo motivo, para se questionar as informações contidas num ranking baseado no valor adicionado, está relacionado com a seguinte questão: para se calcular a efetiva contribuição da escola, deve-se descontar o efeito fora do controle da mesma, como, por exemplo, o *background* familiar dos estudantes. No entanto, ao se fazer isso, aumenta-se a volatilidade do ranking. Por um lado, porque a habilidade do estudante em fazer o teste é medida com erro. Como o evento da prova ocorre durante um dia específico, o estudante pode ter um bom ou mal dia, decorrente de uma doença ou distração, que influencia o resultado do exame. Por outro lado, como apontado por Mcewan *et. al* (2008), (...) *there is sampling variation, since each cohort of students that enters a school is analogous to a random draw from a local population. A school's average performance will thus vary with the specific group of students starting school in any given year. This variance, in turn, depends on the variability of performance in the population from which the school draws its students and the number of students tested.* De fato, Chay *et al.* (2005) mostram que escolas com menos alunos tendem a ter uma maior variância da sua nota, acarretando uma maior volatilidade na sua posição no ranking.¹⁹

18 Para uma experiência recente de política educacional no Brasil no qual bônus são distribuídos em função do valor adicionado da escola, ver Governo do Estado de São Paulo (2009).

19 Usando dados de escolas dos estados americanos da Carolina do Norte e do Sul, Ladd e Walsh (2002) mostram também que esse erro de medida torna o ranking baseado no valor adicionado bastante volátil. Numa análise em educação superior, Yunker (2005) discute a utilidade do con-

As consequências negativas da volatilidade são claras, como apontado por McEwan *et al.* (2008):

While the school rankings may contain some signal of school effectiveness, the added volatility raises concerns. First, ever-shifting conclusions about the 'best' or 'worst' schools may lead parents to sort across schools inefficiently. (...) Second, teachers and school administrators might begin to respond less to incentive programs whose ex-post allocations resemble a lottery. Em outras palavras, a utilidade do ranking torna-se questionável.

Em suma, no que se refere a um ranking baseado no produto, a maior dificuldade está em encontrar uma métrica adequada e consensual para mensurar o produto educacional, principalmente para as IES. No âmbito dos Ensinos Fundamental e Médio existe certo consenso no uso dos resultados nos testes de proficiência para ranquear as escolas. Tanto o ranking baseado no valor absoluto quanto no adicionado tem a sua utilidade, cumprem com objetivos, mas apresentam limitações. O primeiro pode ser “contaminado” pelo resultado da adoção de estratégias oportunistas por parte das instituições de ensino. O principal problema do segundo é a sua volatilidade, que coloca em dúvida a sua utilidade.

3.4 Combinação de Rankings

Existem rankings que são formados a partir de uma combinação de mais de um dos tipos apresentados acima, atribuindo diferentes pesos para cada componente e chegando numa nota final na qual a classificação das instituições é feita.

O problema deste tipo de ranking é que não existe nenhuma base teórica ou empírica que justifique os pesos atribuídos aos seus diferentes componentes. Conforme discutido em Chun (2002), como a classificação final do ranking é sensível às mudanças nos pesos dos seus componentes, torna-se bastante questionável a utilidade do resultado final encontrado.

4 Análise dos dois Principais Rankings Oficiais Brasileiros

A análise na seção anterior apresentou os diferentes problemas e uma visão crítica a respeito dos diferentes tipos de ranking existentes. Com base nesta discussão,

ceito de valor adicionado no curso de contabilidade. Ele mostra que, quando se leva em consideração os intervalos de confiança dos rankings, é possível fazer uma comparação somente entre um pequeno número de instituições.

examinam-se com detalhes, nesta seção, os dois principais rankings oficiais existentes: um para as escolas de Ensino Médio e outro para as IES. O objetivo é chamar a atenção para os seus principais problemas, apontar as suas principais contribuições e sugerir mudanças com o intuito de melhorá-los.

4.1 *Ensino Superior*

O principal problema do ranking do MEC das IES (que é uma combinação de rankings) é a falta de uma justificativa teórica e empírica para a compilação de todas as informações numa só, através da atribuição de pesos para cada um dos seus três componentes, de forma a chegar numa única nota para cada instituição. Isto fica comprovado pela nota técnica explicativa do ranking, que informa que os pesos “foram definidos após diversas reuniões técnicas com especialistas da área de educação superior. Posteriormente o assunto foi discutido na Comissão Nacional de Avaliação da Educação Superior (CONAES), que acabou por referendar os (...) parâmetros”. Como discutido na seção anterior no caso de combinação de rankings, esse é um problema, visto que como a classificação final é sensível a ponderações diferentes, a escolha *ad hoc* torna questionável a utilidade do resultado final encontrado.

No entanto, algum dos componentes do ranking, em separado, tem o potencial de prover informações bastante relevantes para públicos diferentes. Um deles é a média das notas do ENADE somente dos alunos *concluintes*, e não misturada também com a nota dos alunos *ingressantes*. Se essa variável for um bom sinalizador da competência do profissional formado e do seu potencial no mercado profissional, um ranking das instituições baseado somente neste indicador é de extrema relevância para os seus potenciais empregadores. Ter conhecimento deste tipo de ranking também pode ser relevante para os potenciais alunos. Isto porque eles podem querer estar associados com instituições cujas empresas têm forte interesse em contratá-los depois de formados.

As informações contidas num ranking baseado na média das notas do ENADE somente dos alunos *ingressantes* também poderiam ser úteis para os potenciais estudantes. Se essa variável for um bom sinalizador da qualidade do corpo discente, cada aluno pode querer se beneficiar do efeito dos pares (ou *peer effect* em inglês) ao longo do curso e pela sinergia criada por estar em contato com bons alunos.

Outro componente do ranking do MEC, que isolado pode fornecer informações relevantes, é um ranking baseado somente no cálculo do valor adicionado. Por um lado, os estudantes podem também decidir em qual instituição estudar em função de quanto ela agrega para a sua formação. Por outro lado, ele pode ser utilizado pelo

próprio MEC, que fiscaliza a atuação e confere ou retira o credenciamento para o funcionamento da IES.

No entanto, existem pelo menos dois ajustes que poderiam ser feitos no ranking do MEC. O primeiro é referente às informações sobre o valor adicionado da instituição. Primeiro, como o valor adicionado é estimado, tecnicamente o seu valor é uma variável aleatória. Nesse sentido, é importante deixar claro que as diferenças entre os valores adicionados das diferentes instituições podem não ser de fato estatisticamente significativos e elas, apesar de receberem avaliações diferentes, podem na verdade ter a mesma qualidade de ensino. Para contornar este problema, o MEC deveria fornecer os intervalos de confiança dos valores estimados para os valores adicionados de cada instituição.

Por fim, o ranking do MEC é baseado na premissa de que maiores notas no ENADE geram, de alguma maneira, algum benefício para o indivíduo ou a sociedade. Por exemplo, como mencionado acima, maiores nota dos concluintes ou dos ingressantes podem ser um bom sinalizador, respectivamente, da competência do profissional formado ou da qualidade do corpo discente. No entanto, não existe evidência empírica de que comprovem essas possibilidades. Esse ponto é particularmente importante dada a visão, apontada na seção anterior, de que testes de proficiência talvez não mensurem adequadamente o produto educacional de uma IES. Por conseguinte, a busca da comprovação da relevância do ENADE, para o indivíduo ou a sociedade, é fundamental.

4.2 *Ensino Médio*

Como mencionado anteriormente, o ranking oficial das escolas do Ensino Médio é formado a partir das médias das notas obtidas pelos alunos no ENEM.

O principal problema deste ranking é que ele pode, em alguma medida, refletir o comportamento oportunístico por parte das instituições de ensino. Como mencionado anteriormente, é possível que as instituições expulsem os piores alunos de forma a que eles não representem a escola no ENEM, levando ao aumento da sua média e classificação no ranking. Por conseguinte, as escolas preocupadas em recuperar alunos com dificuldades seriam piores ranqueadas.

Caso os pais/alunos escolham as suas escolas com base neste ranking, isto pode levar a uma segregação no mercado. Isto porque as escolas melhores ranqueadas teriam um excesso de demanda, selecionariam cada vez mais os melhores alunos dentro do pool maior disponível de candidatos e, como resultado, manteriam ou melhorariam suas posições no ranking. O contrário verificar-se-ia com as escolas relativamente

piores. O problema é que estas últimas podem ser as mais eficientes na produção do serviço educacional.

Mesmo sendo possivelmente influenciado por comportamentos oportunistas, o ranking das escolas tem o potencial de prover informações relevantes para diferentes públicos. Por um lado, se a nota do ENEM sinalizar a futura competência do estudante do Ensino Médio que segue para o mercado de trabalho, o ranking é relevante para os potenciais empregadores. Por outro lado, se a nota do ENEM sinalizar a qualidade do corpo discente da escola, o ranking é relevante para os pais no momento de escolher a escola dos seus filhos, devido ao efeito dos pares. Por fim, se a nota do ENEM sinalizar a capacidade do aluno de completar com sucesso a graduação, os pais/alunos teriam interesse em utilizar o ranking.

No entanto, não existe evidência empírica que comprove que o ENEM tenha a capacidade de sinalizar a competência do estudante no Ensino Médio, no mercado de trabalho ou em estudos avançados. Estudos sobre a efetiva capacidade do resultado no ENEM de prever impactos significativos para o indivíduo ou a sociedade são necessários. Assim como no caso do ENADE, o risco é que se foque num alvo diferente da qualidade da educação, ao se dar muita importância aos resultados do ENEM.²⁰

Por fim, para contornar o problema relacionado a comportamentos oportunistas mencionado acima, o MEC deveria divulgar, simultaneamente ao ranking existente, outro baseado no valor adicionado. Neste cálculo, deveria levar em consideração a razão entre o número de concluintes e ingressantes no curso no cálculo do mesmo, como é feito no caso das IES.

5 Conclusão

A discussão no trabalho sinaliza que qualquer tipo de ranking em educação apresenta problemas e são passíveis de críticas. Dado que nenhum ranking é perfeito, o importante é o entendimento de que tipo de informação cada um fornece.

Foram examinados, com detalhes, os dois rankings oficiais das IES e do Ensino Médio. Um problema em comum é que não existem evidências de que os resultados das provas (ENEM e ENADE) tenham a capacidade de sinalizar a competência do

20 Esse tópico, nos dias atuais, é particularmente importante porque o MEC já comunicou que vai mudar o formato do ENEM; várias universidades federais já se comprometeram a utilizar o resultado desse Exame pelo menos como parte dos seus processos seletivos. Para maiores detalhes sobre as mudanças no ENEM, ver INEP (2009) (http://portal.mec.gov.br/index.php?option=com_content&view=article&id=13318&Itemid=31).

estudante, seja no mercado de trabalho ou em estudos avançados. Como consequência, é possível que esses resultados não mensurem adequadamente o produto educacional.

No caso do ranking das IES, outro problema é o fato dele ser formado por vários componentes que são compilados num único indicador. Devido à falta de justificativa para os pesos atribuídos aos diferentes componentes e o fato da classificação final ser sensível a ponderações diferentes, a escolha *ad hoc* torna questionável a utilidade do resultado final encontrado. A utilização de cada uma dos seus componentes em separado, no entanto, pode trazer informações relevantes para diferentes *stakeholders*.

No caso do ranking das escolas de Ensino Médio, ele pode refletir o comportamento oportunístico das instituições, beneficiando aquelas com menor preocupação em recuperar os alunos com maiores dificuldades de aprendizado em detrimento daquelas com políticas mais inclusivas. O problema é que estas últimas podem ser as mais eficientes na produção do serviço educacional.

Neste sentido, é interessante notar que o MEC tem uma política totalmente distinta para o ranking das IES vis-à-vis às do Ensino Médio. Para as primeiras, além de fornecer informações sobre as notas médias do ENADE por instituição, como faz no caso do ENEM para as escolas, ele estima um indicador da contribuição da instituição para o aprendizado do aluno, ou o valor adicionado. Adicionalmente, ele se preocupa em evitar que as IES que porventura expulsem os seus piores alunos de forma a aumentar a média da instituição sejam beneficiadas no ranking. É surpreendente que a mesma preocupação não exista no caso do ranking do Ensino Médio.

Dado os impactos dos rankings oficiais, seria importante que ajustes fossem feitos de forma a aprimorá-los. Com isto, os agentes econômicos poderiam tomar suas decisões de qual instituição de ensino escolher de forma melhor embasada, com impactos positivos sobre a alocação dos recursos na economia.

Referências

- ANDRADE, E.; MOITA, R.; SILVA, C. (2009) *A escolha da faculdade pelo aluno: estimação da demanda e precificação dos atributos*; Mimeografado.
- BISHOP, J. (1991) Achievement, test score, and relative wage; em *Workers and Their Wages*; Washington, DC: The AEI Press.
- CHAY, K.; MCEWAN, P.; URQUIOLA, M. (2005) The central role of noise in evaluating interventions that use test scores to rank schools. *American Economic Review*, Vol. 95, Issue 4, September.
- CHUN, M. (2002) Looking where the light is better: a review of the literature on assessing higher education quality. *Peer Review*, Winter/Spring.
- CLARKE, M. (2002) Quantifying quality: what can the U.S. news and world report rankings tell us about the quality of higher education? *Education Policy Analysis Archives*, Vol. 10, Issue 16, March.
- CUNHA, F.; HECKMAN, J.; LOCHNER, L.; MASTEROV, D. (2005) *Interpreting the evidence on life cycle skill formation*. NBER Working Paper #11331.
- GOVERNO DO ESTADO DE SÃO PAULO (2009) *Programa de qualidade da escola – Nota técnica*; Mimeografado.
- GUIA DO ESTUDANTE (2008). *As melhores universidades do país*. Editora Abril.
- HANUSHEK, E. (1979) Conceptual and empirical issues in the estimation of educational production functions. *The Journal of Human Resources*, Vol. 14, Issue 3, Summer.
- HANUSHEK, E. (2003) The failure of input-based schooling policies. *The Economic Journal*, Vol. 113, Issue 485, February.
- HANUSHEK, E. WOESSMANN, L. (2008) The role of cognitive skills in economic development. *Journal of Economic Literature*, Vol. 46, Issue 3, September.
- INEP (2008) *Cálculo do conceito preliminar de cursos de graduação – Nota Técnica*. http://enade.inep.gov.br/enadeResultado/pdfs/Procedimentos_Metodologicos_Calculo.pdf; visita ao site no dia 26/05/2009.
- INEP (2009); ENEM; http://portal.mec.gov.br/index.php?option=com_content&view=article&id=13318&Itemid=310; visita ao site no dia 22/06/2009.
- LADD, H.; WALSH, W. (2002) Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, Vol. 21, Issue 1, February.
- MCEWAN, P.; URQUIOLA, M. VEGAS, E. (2008) School choice, stratification, and information on school performance: lessons from Chile. *Economia: Journal of the Latin American and Caribbean Economic Association*, Vol. 8, Issue 2c, Spring.

- MENEZES-FILHO, N. (2007) *Os determinantes do desempenho escolar do Brasil*. Mimeografado.
- MONKS, J.; EHRENBERG, R. (1999) *The impact of U.S. news & world report college rankings on admissions outcomes and pricing policies at selective private institutions*, NBER, Working Paper # 7227.
- MURPHY, K.; PELTZMAN, S. (2004) School performance and the youth labor market. *Journal of Labor Economics*, Vol. 22, Issue 2, April.
- O ESTADO DE SÃO PAULO (2008) *Particular quer fim da divulgação de avaliações*. Publicado na edição de 5/12/2008.
- SCHWARTZMAN, S. (2004) *The national assessment of courses in Brazil*. Public Policy of Academic Quality Research Program; March.
- YUNKER, J. (2005) The dubious utility of the value-added concept in higher education: the case of accounting. *Economics of Education Review*, Vol. 24, Issue 3, June.