

A qualidade do ranking das escolas de ensino médio baseado no ENEM é questionável

Eduardo Andrade

Economista

(APEX Capital Ltda.)

Endereço: R. Gomes de Carvalho 1069/143 – Vila Olímpia, São Paulo/SP – Brasil

CEP: 04547-004 – E-mail: eduardo.andrade@apexcapital.com.br

Ivan Soida

Professor (Escola Superior de Propaganda e Marketing)

Endereço: Rua Doutor Álvaro Alvim, 123 - Vila Mariana, São Paulo/SP – Brasil

CEP: 04018-010 – E-mail: soida.ivan@gmail.com

Recebido em 22 de novembro de 2012. Aceito em 08 de agosto de 2014.

Resumo

Este trabalho questiona a qualidade do ranking das escolas do ensino médio baseado no ENEM. Escolas menores tendem a apresentar desempenho: (i) mais extremado; (ii) com flutuações de curto prazo mais significativas; e (iii) com maior peso de “fatores não persistentes” na decomposição da variância da sua nota. Adicionalmente, o ranking tem um componente transitório importante, que tende a desaparecer na edição seguinte. Por fim, a probabilidade de uma escola manter, no futuro, uma determinada posição no ranking é relativamente baixa. Em suma, a qualidade das informações contida no ranking é baixa.

Palavras Chave

ENEM. Educação. Ensino Médio. Rankings.

Abstract

The findings in this paper question the quality of the high schools’ ranking based on ENEM. Small schools’ test scores tend to: (i) be more extreme; (ii) have larger short run fluctuations; and (iii) have a greater weight of “non-persistent factors” in the decomposition of their scores’ variance. Moreover, the ranking has an important temporary component that tends to disappear in the next edition. Finally, the probability that one school maintains, in the future, a given position in the ranking is small. In sum, the quality of the information provided by the ranking is low.

Keywords

ENEM. Education. High School. Rankings.

JEL Classification

I21. I2. I20. I22. I23. I28. I29

1. Introdução

O ENEM (Exame Nacional do Ensino Médio) é uma prova de abrangência nacional, aplicada anualmente pelo Inep (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), órgão vinculado ao MEC (Ministério da Educação). Apesar de o exame ter sido criado, em 1998, com a finalidade principal de avaliar o perfil de saída dos concluintes do ensino médio no país, ele gradualmente incorporou uma série de outros usos: como critério parcial ou exclusivo para ingresso à ampla maioria das universidades federais,¹ como principal quesito de acesso a bolsas de estudo e crédito estudantil financiado pelo governo² e, recentemente, como um certificador oficial de conclusão do ensino médio no país.³

Entretanto, uma dessas finalidades gerou maior repercussão na sociedade, graças, sobretudo, à ampla divulgação nos principais canais da mídia: a utilização dos resultados do ENEM para estabelecer um ranking nacional de escolas do ensino médio. Anualmente, 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. O ranking do ENEM é o único indicador de desempenho, de abrangência nacional, para as escolas de ensino médio, e tem sido divulgado pelos órgãos oficiais a partir de 2004.

Existe evidência de que rankings em educação afetam o mercado. Por exemplo, no caso brasileiro, Andrade *et al.* (2009) mostram que a posição de uma instituição de ensino superior no ranking do Guia de Estudante afeta de forma significativa o seu *market share*. Monks e Ehrenberg (1999) avaliam o impacto do ranking da revista U.S News and World Report, o mais tradicional do mercado americano, sobre as universidades dos EUA. Eles concluem que uma pior posição no ranking é prejudicial para a universidade em vários sentidos – uma menor parcela dos alunos aceitos pela universidade decide se

¹ Houve grande investimento do MEC nos últimos anos no sentido de estabelecer o ENEM como o principal e, eventualmente, o único critério para ingresso nas disputadas instituições federais de ensino superior, através do Sistema de Seleção Unificada (Sisu). A seleção é feita com base na nota obtida pelo candidato no ENEM. Outros vestibulares tradicionais, como a FUVEST, também já adotaram a nota no ENEM como critério parcial para seus processos seletivos. Para maiores detalhes sobre o Sisu, ver <http://sisu.mec.gov.br/#/principal.php>.

² Sobre o uso do ENEM para a concessão de bolsa de estudo, ver <http://siteprouni.mec.gov.br/>.

³ Sobre o uso do ENEM para esta certificação, ver <http://encceja.inep.gov.br/certificacao>.

matricular, a qualidade das novas turmas cai e o preço líquido pago pelo aluno é menor.

A evidência anedótica sugere que os efeitos do ranking do ENEM 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.⁴ Camargo *et al.* (2011) sugerem que a divulgação dos resultados do ENEM pode ter um impacto sobre o desempenho futuro dos alunos.

Dados os seus potenciais impactos e a importância do ranking do ENEM, é importante examinar a qualidade do ranking e se os *stakeholders* (pais/estudantes, diretores de escolas, *policymakers* etc.) das escolas deveriam levar em consideração as informações contidas nele no momento de tomar as suas decisões. É exatamente este o objetivo deste trabalho. Testar, de diferentes maneiras, a qualidade do ranking do ENEM.

Inicialmente, seguindo a metodologia desenvolvida em Kane e Staiger (2002), mostramos que o tamanho de uma escola, mensurado pelo número de alunos que realizam o ENEM, afeta o desempenho da mesma no ranking. Escolas menores estão mais associadas a *scores* extremos (baixos ou altos). Adicionalmente, as flutuações no desempenho de um ano para outro são mais significativas nas escolas de menor porte. Esses resultados, qualitativamente iguais aos verificados em Kane e Staiger (2002), não surpreendem, porque não existem motivos teóricos para que sejam diferentes. No entanto, eles são uma importante evidência que questionam a qualidade do ranking baseado no ENEM.

A novidade da análise neste trabalho em relação à análise de Kane e Staiger (2002) está na comparação entre o desempenho das escolas privadas e públicas. As primeiras apresentam uma distribuição de notas menos homogênea, talvez por serem escolas na média de menor porte, e com rendimentos superiores.

⁴ Ver artigos: “Escolas no topo do ranking do Enem vão reajustar mensalidade em até 10%” no G1 de 13/11/2009, “Mensalidade das melhores escolas da capital sobe o dobro da inflação” no OESP de 3/12/2009 e “Escolas particulares: Ponto na prova ‘custa’ de R\$ 16 a R\$ 52” no OESP de 12/09/2011.

Por fim, fatores não persistentes, ou seja, aqueles geralmente fora do controle da escola (e que não podem ser repetidos regularmente a cada medição ou exame do ENEM) podem explicar algo em torno de 60% da variância dos resultados de escolas pequenas vis-à-vis 16% para escolas grandes, no caso de escolas públicas (e, no caso de escolas privadas, 37% vs. 9%).

Em seguida, utilizamos a metodologia desenvolvida em Dichev (1999) para testar a previsibilidade do ranking baseado no ENEM. O resultado encontrado sugere que o ranking exibe uma autocorrelação significativa e negativa nas mudanças das notas. Isto sugere que o ranking tem um componente transitório importante, que tende a desaparecer na edição seguinte. Por conseguinte, o ranking presente não seria uma boa aproximação de rankings futuros.

Por fim, utilizando matrizes markovianas, avaliamos a proporção de escolas entre as 10% melhores (ou qualquer outra posição relativa no ranking) que permaneceram nessa posição ao longo do tempo. Essa análise foi feita tanto em escala nacional quanto para as escolas do município de São Paulo. Os resultados sugerem que existe uma mudança importante nas posições do ranking de um ano para outro. Por exemplo, no município de São Paulo, caso pais/alunos tivessem elegido uma escola entre as 10% melhores no ranking do ENEM em 2006 para efetuar uma matrícula, havia uma probabilidade de somente 28% de que essa escola permanecesse nessa mesma elite em 2009 (quando o estudante concluiu o ensino médio).

A combinação dos resultados obtidos neste trabalho coloca em dúvida a utilização dos resultados do ENEM como uma boa métrica para auferir a qualidade das escolas do ensino médio. A qualidade das informações nele contida em determinado ano parece de pouca valia para guiar a escolha futura de, por exemplo, pais/alunos, tendo em vista a duração do ensino médio (ou seja, ao menos três anos). A influência do ranking do ENEM sobre os *stakeholders* das escolas pode conduzir a decisões inadequadas, gerando alocação ineficiente de recursos, pois ele é pautado por informações de qualidade duvidosa. Por exemplo, uma escola momentaneamente bem posicionada no ranking pode gerar forte demanda por seus serviços e captar grande volume de recursos, mas sua qualidade real pode ser bem inferior ao divulgado.

De acordo com o nosso conhecimento, este é o primeiro trabalho que examina a qualidade do ranking das escolas médias no Brasil baseado no ENEM. Este trabalho está dividido em cinco seções, incluindo esta introdução. Na seção seguinte, descrevemos as principais características do ranking do ENEM e apresentamos a base de dados utilizada neste trabalho. Na seção 3, analisamos a relação existente entre o tamanho da escola e o seu desempenho no ENEM. Na seção 4, testamos a previsibilidade e a estabilidade do ranking. A última seção conclui.

2. Ranking do ENEM e base de dados

O ENEM é uma avaliação de adesão voluntária, disponível para todos os concluintes do ensino médio ou pessoas já formadas neste nível. Em sua primeira aplicação, em 1998, 157.221 pessoas realizaram o exame. Desde então, o número de participantes tem aumentado gradativamente e atingiu o patamar de 5.366.780 em sua edição mais recente, em 2011.⁵ Esse significativo aumento recente (em 2009, eram pouco mais de 4 milhões os inscritos no ENEM)⁶ explica-se, em grande medida, pela inclusão do ENEM nos processos seletivos das universidades públicas federais.

Para os objetivos deste trabalho, a mais importante modificação introduzida pelo MEC ocorreu em 2004, quando ele passou a divulgar as médias, obtidas pelos alunos de cada escola, na prova do ENEM. Esse foi o ponto de partida para a criação do ranking das escolas do ensino médio baseado nos resultados do ENEM, em que a posição de uma escola é determinada exclusivamente pelo desempenho de seus alunos na prova.

O ENEM, atualmente, é composto por quatro provas objetivas (Humanidades, Ciências, Matemática e Linguagens), cada uma com 45 questões objetivas, e uma redação. Um candidato é considerado elegível para fins de cálculo de seu desempenho (e do desempenho da escola na qual ele declarou ter concluído o ensino médio), caso

⁵ A partir de 2009, o MEC optou pela metodologia de Teoria de Resposta ao Item (TRI) em substituição à Teoria Clássica de Resposta (TCR), para permitir a comparabilidade intertemporal dos resultados dali em diante.

⁶ A fonte da informação dos números de inscritos no ENEM é o Inep.

participe das quatro provas objetivas e da prova de redação naquele ano. A nota final de um aluno elegível é igual a uma média ponderada, na qual cada prova objetiva tem o mesmo peso (12,5% cada) e a prova de redação tem peso igual ao conjunto das demais (50%).

O Inep divulga três notas para cada escola: média simples dos alunos da instituição nas provas objetivas, média simples dos alunos da instituição na prova de redação e média aritmética simples entre esses dois *scores* para cada escola.⁷ O resultado com mais ampla divulgação na mídia é o último, que incorpora a combinação das notas das provas objetivas e de redação.

Como a prova é voluntária, é possível que poucos alunos de uma determinada escola a realizem. Os resultados das escolas sub-representadas, cujo número de alunos seja muito pequeno e pouco representativo do universo de alunos da escola, no entanto, não são computados, e o resultado das mesmas não aparece no ranking. Isto porque o MEC frequentemente adota dois critérios mínimos de elegibilidade para uma escola entrar no ranking das escolas do ensino médio baseado no ENEM, com sua consequente publicação. Quais sejam, escolas com menos de 10 alunos e/ou com menos de 2% (dos alunos matriculados) participantes da prova são excluídas das listagens daquela edição do exame. Em 2006, foi aplicado um terceiro mecanismo, que era um fator de correção (um ajuste estatístico) da nota, em função do percentual de alunos da escola que realizou a prova, a fim de neutralizar o viés de participação.⁸

Em 2010, último ranking disponível, o MEC publicou quatro rankings distintos, em que as escolas eram agrupadas por bandas de taxa de participação: 2% a 25%, 25% a 50%, 50% a 75%, 75% a 100% dos alunos matriculados. Essa forma de divulgação do resultado tenta evitar um potencial problema do ranking: o fato de que ele pode, em alguma medida, refletir o comportamento oportunístico por parte das instituições de ensino. Em outras palavras, é possível que instituições acionem mecanismos diversos para expulsar seus piores alunos, de forma que eles não representem a escola no ENEM,

⁷ As notas individuais não são divulgadas.

⁸ Em 2006, o Inep, ao perceber que a performance da escola no ENEM era função do número de alunos prestando o exame, testou várias especificações (modelos) de correção. Finalmente, chegou a um modelo quadrático côncavo, que foi usado para estimar as notas efetivas. Na ocasião, quatro notas (objetiva com e sem uso do corretor e viés, geral com e sem uso do corretor de viés) foram divulgadas. De todo modo, segundo o Inep, o efeito desse corretor foi marginal e insignificante para a maioria das escolas.

inflando artificialmente a média da escola e, por conseguinte, alcançando uma melhor classificação no ranking. Em paralelo, as escolas preocupadas em recuperar alunos com dificuldades seriam piores ranqueadas, mesmo que em tese possam ser mais eficientes do ponto de vista educacional.⁹

A base de dados para todas as análises feitas neste trabalho é a mesma: os resultados das edições do ENEM de 2006 a 2010, que são divulgados pelo Inep. Desta base de dados são extraídas as seguintes informações, para cada escola do ensino médio regular,¹⁰ pública ou privada: número de alunos participantes da prova objetiva, número de participantes da prova de redação, média geral na prova objetiva, média geral na prova de redação e média geral final. Esta última é calculada como a média simples entre as médias “geral da prova objetiva” e “geral na prova de redação”.

Um procedimento adotado na manipulação dos dados do ano de 2010 merece atenção. Diferentemente dos demais anos, não constava o código da escola na listagem oficial disponibilizada pelo Inep para o ano de 2010. A alternativa encontrada foi criar um rótulo indexador (um “label”) formado pela combinação (“nome da escola” + “unidade da federação”) para cada escola preexistente nas bases de dados de 2006 a 2009, e comparar esse indexador com “labels” análogos criados para a listagem de 2010. Somente nos casos de identidade perfeita entre os dois “labels” os dados referentes ao ano de 2010 foram incorporados à base de dados. Nos demais casos, foram descartados.¹¹

Esse procedimento reconheceu 11.269 escolas em nível nacional, permitindo um aproveitamento, ao menos parcial, dos dados de 2010. Porém, para outras 9.929 escolas, isso não foi possível; esse último conjunto de dados foi sumariamente desconsiderado naqueles ensaios que envolveram os scores de 2010.

Desta forma, os números de escolas presentes nas séries temporais para o triênio 2008-2010 e biênio 2008-2009 são, respectivamente, iguais a 8.511 e 15.171. Adicionalmente, todas as análises que utili-

⁹ Para maiores detalhes sobre esta discussão, ver Andrade (2011).

¹⁰ Não foram utilizadas as informações relativas às escolas em modalidades minoritárias, como EJA (Ensino de Jovens e Adultos), EMT (Ensino Médio Técnico-Profissionalizante) e categorias mistas, quando essa discriminação estava disponível.

¹¹ Consideramos mínima a possibilidade de ocorrência de escolas diferentes com o mesmo “label”.

zam os dados de 2006 e 2007 são restritas para as escolas do estado de São Paulo.¹² Desta forma, os números de escolas presentes nas séries temporais para o quinquênio 2006-2010, quadriênio 2006-2009, quadriênio 2007-2010 e triênio 2007-2009 são, respectivamente, iguais a 1.753, 3.821, 1.823 e 3.441.

Por fim, para a análise da decomposição da variância da nota da escola, explicada na próxima seção, foi utilizada uma informação adicional, também extraída da base de dados do ENEM tornada pública pelo Inep. A partir dos microdados do ENEM para o ano de 2006, foi calculada a variância das notas do ENEM (ou variância intraescolar), obviamente utilizando as notas de todos os alunos que compuseram a média da escola. Esse cálculo foi feito para cada escola do estado de São Paulo.

3. Tamanho da escola e desempenho no ENEM

Esta seção examina potenciais problemas ocasionados pelo uso dos resultados do ENEM para ranquear as escolas do ensino médio. Esses problemas estão relacionados à imprecisão da medida utilizada, a choques fora do controle da escola e à interferência de fatores transitórios, que provocam flutuações de curto prazo, de natureza estatística.

Como resultado, questiona-se a qualidade desta medida para comparar o desempenho tanto entre escolas quanto de uma mesma instituição ao longo do tempo. Esse efeito é ainda mais significativo para as escolas com menor número de alunos. Seguindo a metodologia utilizada em Kane e Staiger (2002), inicialmente, mostramos como a nota da escola, num determinado ano, e a mudança da nota, entre anos, estão correlacionadas ao tamanho da escola. Em seguida, examinamos se a melhora (ou a piora) do desempenho das escolas em um determinado ano tende a ser revertida num ano seguinte. Por fim, decompomos a variação da nota da escola no ENEM para quantificarmos a importância relativa de fatores persistentes e não persistentes.

¹² No momento da elaboração do trabalho, para os anos 2006 e 2007, os autores tiveram acesso somente aos dados para o Estado de São Paulo.

Antes de apresentar os resultados, é importante informar que as análises nesta seção utilizam as notas médias dos alunos de cada escola na parte objetiva da prova do ENEM. Isto se deve à maior objetividade implícita ao processo de correção nesse tipo de avaliação, em comparação à prova de redação, que tende a incorporar algum grau de subjetividade pelos corretores. Ou seja, não utiliza o critério oficial de ranqueamento adotado pelo MEC, que combina a média entre a prova objetiva e a redação, dando pesos iguais (50%) para ambas as provas. A justificativa é que um ranking, em tese, seja de melhor qualidade quando um critério mais objetivo de avaliação das escolas é utilizado. Por conseguinte, caso o ranking utilizando somente o resultado das provas objetivas não seja de boa qualidade, dificilmente o ranking baseado também no resultado da redação o seria.

O Gráfico 1 plota, em gráficos de dispersão, os resultados do ENEM 2010, média da escola na prova objetiva, contra o número de alunos que realizaram a prova de cada instituição.¹³ Os Gráficos (1a), (1b) e (1c) mostram as dispersões obtidas para escolas do Brasil – respectivamente, todas as escolas (19.414 escolas), somente escolas privadas (5.167 escolas) e somente escolas públicas (14.247 escolas).

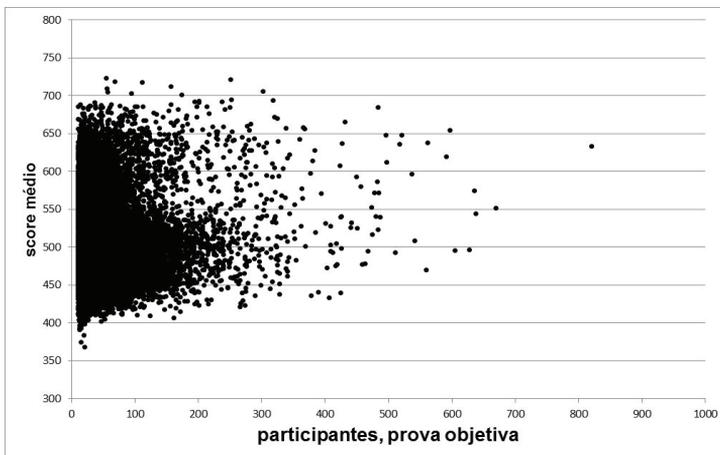


Gráfico 1(a). Médias por escola, nacional, prova objetiva.

¹³ Como mencionado na seção anterior, o MEC somente divulga a nota da instituição que teve ao menos 10 alunos participantes de cada prova e com mais de 2% dos alunos matriculados realizando a prova, no ano em questão. Seguimos esse mesmo critério neste trabalho.

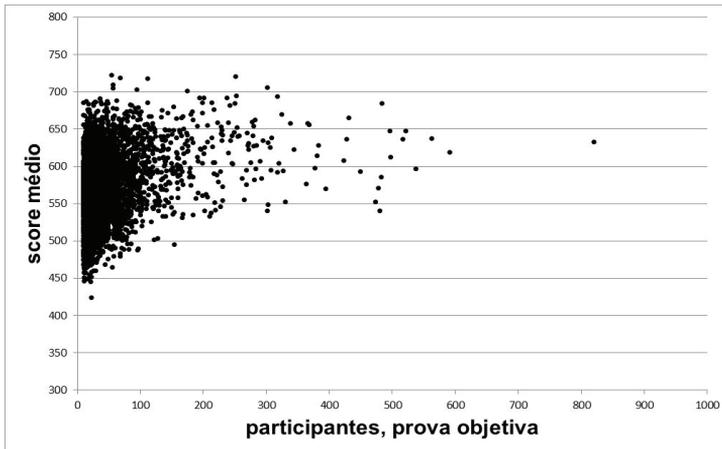


Gráfico 1(b). Médias por escolas privadas, nacional, prova objetiva.

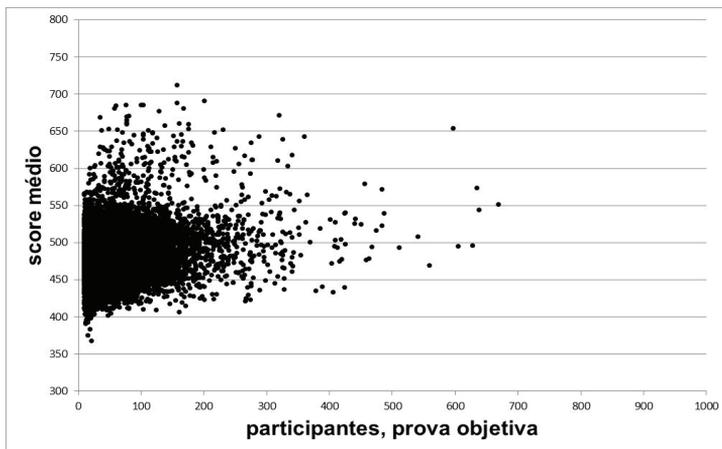


Gráfico 1(c). Médias por escolas públicas, nacional, prova objetiva.

Fonte: Inep. Dados trabalhados pelos autores.

Um exame destes gráficos sugere que, em linhas gerais, existe uma maior variabilidade entre as escolas com um menor número de alunos. Em outras palavras, escolas de pequeno porte estão associadas a distribuições mais amplas e a scores extremos (baixos e altos), enquanto as escolas de grande porte exibem scores mais centralizados. Como apontado por Kane e Staiger (2002), em escolas com

um grupo reduzido de alunos, “even a few particularly talented or rambunctious youth can have a large impact on the grade-level score for a school from year to year”.

Uma comparação entre o desempenho das escolas privadas e públicas (Gráficos (1b) e (1c)) revela o seguinte. As primeiras são em menor número, com as suas notas distribuídas de modo mais heterogêneo e com rendimentos médios superiores em relação às escolas públicas (privadas com nota média de 587,44 vs. 490,77 nas públicas). Talvez a sua distribuição mais heterogênea seja reflexo do menor tamanho médio das escolas privadas (41,22 alunos por instituição, contra 54,29 na pública). Essa comparação entre escolas privadas e públicas é a novidade em relação ao trabalho de Kane e Staiger (2002).

O Gráfico 2 plota, em gráficos de dispersão, a mudança da nota na prova objetiva do ENEM entre os anos de 2008 e 2009 contra o número de alunos que realizaram a prova de cada instituição,¹⁴ para todas as escolas do Brasil.

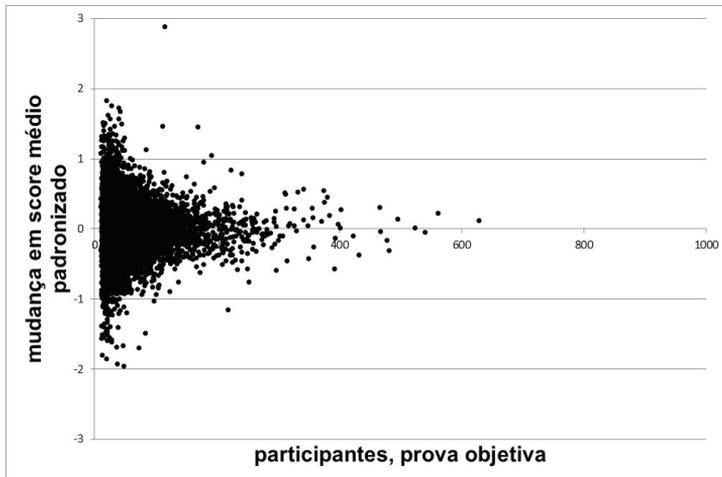


Gráfico 2: mudança em médias padronizadas da prova objetiva por escolas, nacional (2008/2009)

Fonte: Inep. Dados trabalhados pelos autores.

¹⁴ Tomamos as listagens do ENEM referentes às médias por escola de 2008 e 2009 por serem as bases mais completas disponíveis. As médias foram padronizadas para a análise, ou seja, foi subtraída de cada valor a média do respectivo ano e o resultado foi dividido pelo desvio-padrão anual.

O diferencial de desempenho entre as escolas de menor porte vis-à-vis as de maior porte é mais nítido agora. As flutuações de curto prazo das notas do ENEM são muito mais aparentes nas escolas de menor porte. Existe uma volatilidade na medição do desempenho das escolas pequenas, com o desempenho variando de forma mais significativa de um ano para outro. As escolas de grande porte, por outro lado, tendem a exibir maior estabilidade nessas medidas de desempenho, com mudanças de menor magnitude entre anos.

Como aponta Kane e Staiger (2002), existem várias explicações para esta flutuação de curto prazo no desempenho das escolas. Uma delas é a variação amostral. Por exemplo, pode existir uma coorte de desempenho excepcionalmente bom (o que os professores informalmente chamariam de “boa safra”) num ano, seguido por outra excepcionalmente ruim no ano seguinte, impactando na média final atingida por aquela instituição no ENEM ao longo dos anos. Outros fatores podem explicar estas flutuações temporárias, como um surto de gripe forte numa determinada região, um canteiro de obras em operação no dia de aplicação do exame ou uma maior coincidência entre o material enfatizado num período de revisão para a prova e o conteúdo do exame.

Seria interessante testar a importância da variação amostral ou de outros choques temporários sobre o desempenho das escolas. Uma forma de realizar isto é examinar se a nota da escola num determinado ano tende a ser revertida no ano seguinte. É exatamente isto que realizamos em seguida.

O Gráfico 3 plota o coeficiente de correlação entre as notas da prova objetiva do ENEM de um até três anos à frente, entre os anos de 2006 e 2009, para as escolas *públicas* do Estado de São Paulo. Os Gráficos (3a), (3b) e (3c) mostram os resultados, respectivamente, para o 1º. (494 escolas), 3º. (495 escolas) e 5º. (495 escolas) quintis.¹⁵

Em cada gráfico, o primeiro ponto mostra o coeficiente de correlação entre as notas do ENEM de 2006 com elas mesmas, ou seja, com “zero” anos de diferença entre as notas (variável no eixo horizontal). Obviamente, esse coeficiente é igual a 1, que aparece no eixo verti-

¹⁵ As escolas foram distribuídas dentro dos quintis em função do número médio de alunos que prestaram a prova do ENEM, de 2006 a 2009. As notas foram padronizadas em cada ano.

cal. O ponto seguinte registra o coeficiente de correlação das notas do ENEM de 2006 com aquelas verificadas no ano de 2007, ou seja, com um ano de diferença entre as notas, e assim sucessivamente.

Do ponto de vista qualitativo, a análise dos três gráficos (3a, 3b e 3c) mostra um mesmo comportamento. Verifica-se uma queda significativa na correlação com somente um ano de diferença entre as notas. Adicionalmente, passado o primeiro ano, ocorrem quedas graduais de correlação ao longo do tempo.

Kane e Staiger (2002) explicam esse comportamento. Eles sugerem que, sobre as notas de testes de proficiência como o ENEM, incidem duas ordens de fatores. Por um lado, existem os “fatores não persistentes”. Esses estão geralmente fora do controle da instituição e, portanto, não podem ser repetidos regularmente a cada medição. Podem ter natureza amostral, como uma coorte de alunos de rendimento muito acima, ou muito abaixo do comum, ou natureza eventual, como uma condição climática extrema, ou uma escavadeira em operação no dia da prova, afetando o rendimento dos alunos. Como dificilmente se repetem de um ano para outro, são altamente voláteis, tendo impacto apenas no curto prazo e, portanto, tendem a sofrer uma rápida atenuação temporal.

Pode-se argumentar que as escolas mal ranqueadas reajam investindo mais para melhorar o desempenho dos seus alunos. Em contraste, as escolas bem ranqueadas talvez invistam menos e piorem a sua posição no ranking. Dessa forma, as variações do componente não persistente podem vir do ruído estatístico e não de melhorias de qualidade. Apesar de em tese ser possível, esse tipo de reação das escolas gera resultados no longo prazo, pois é razoável assumir que exista um *delay* entre o investimento feito e o ganho de desempenho. De toda forma, é pouco provável que esse argumento seja válido para escolas públicas, visto que o desempenho das mesmas no ENEM gera pouco ou nenhuma repercussão em termos de ação/investimento da parte dos seus gestores.

Por outro lado, existem os “fatores persistentes”. Esses estão geralmente sob o controle da instituição, como a composição do corpo docente, o currículo adotado e as condições de infraestrutura. Essas são relativamente estáveis de uma medição para outra e sofrem declínio temporal lento. Por exemplo, a troca de professores da equipe, as

alterações curriculares e a depreciação natural da infraestrutura escolar são fenômenos que, geralmente, ocorrem de modo gradual. Portanto, essa baixa correlação entre as notas do ENEM com somente um ano de diferença provavelmente reflete o efeito, volátil, destes “fatores não persistentes”. Já as mudanças graduais, após o primeiro ano, devem refletir as pequenas mudanças associadas aos “fatores persistentes”.

É interessante comparar a trajetória do coeficiente de correlação em função do porte das escolas. Na comparação entre quintis (Gráficos (3a)-(3c)), observa-se uma relação negativa entre a queda da correlação com somente um ano de diferença entre as notas e o tamanho das escolas. A correlação passou de 0,35 para 0,64, e para 0,76, respectivamente, do primeiro para o terceiro, e para o quinto quintil. Esses resultados sugerem que escolas com um menor número de alunos apresentam comportamento significativamente mais volátil, indicando que os “fatores não persistentes” têm maior importância relativa aqui do que têm em escolas de maior porte. Usando essa nova abordagem, esses novos resultados corroboram aquele já encontrado anteriormente: as flutuações de curto prazo das notas do ENEM tendem a ser mais significativas nas escolas de menor porte.

O Gráfico 4 plota o coeficiente de correlação entre as notas do ENEM de um até três anos à frente, entre os anos de 2006 e 2009, agora para as escolas *privadas* do Estado de São Paulo. Os Gráficos (4a), (4b) e (4c) mostram os resultados, respectivamente, para o 1º. (145 escolas), 3º. (145 escolas) e 5º. (145 escolas) quintis.¹⁶

Todos os resultados encontrados na análise para as escolas públicas são reeditados agora para as escolas privadas, mesmo essas apresentando um número pequeno de observações em cada quintil: há uma queda significativa na correlação com somente um ano de diferença entre as medições; passado o primeiro ano, ocorrem quedas graduais da correlação ao longo do tempo; e, por fim, relação negativa entre a queda da correlação com somente um ano de diferença entre as notas e o tamanho das escolas.

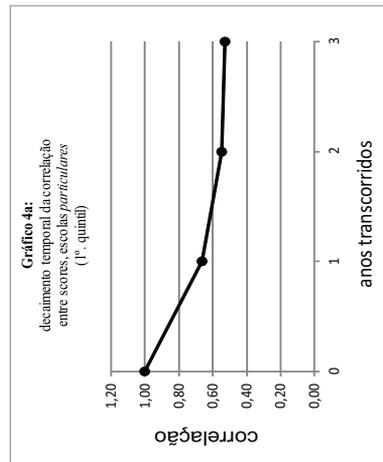
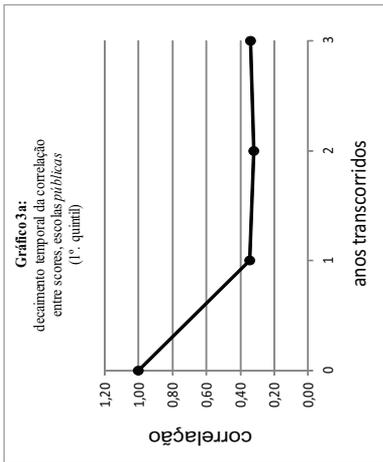
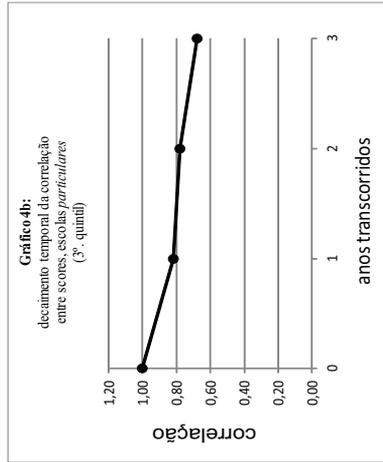
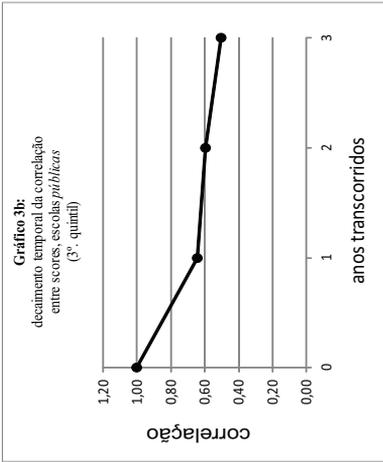
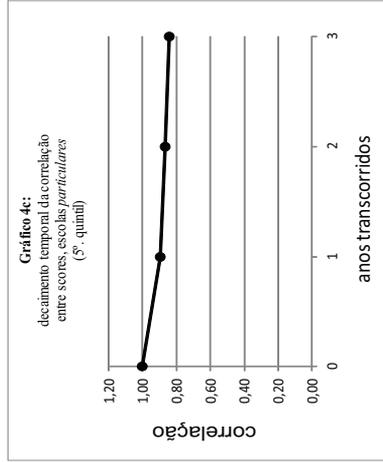
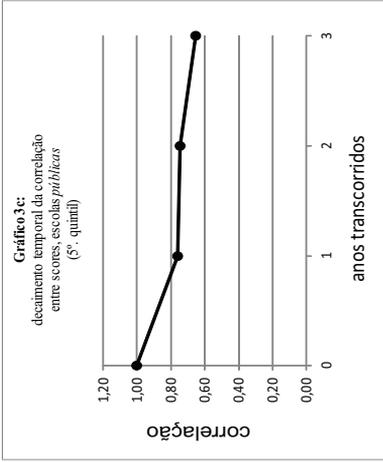
¹⁶ Assim como no caso das escolas públicas, as escolas foram distribuídas dentro dos quintis em função do número médio de alunos que prestaram a prova do ENEM nos anos de 2006 a 2009. As notas foram padronizadas em cada ano.

O interessante agora é a comparação entre os dois universos, escolas públicas e privadas. Independentemente do quintil escolhido, a queda do coeficiente de correlação com somente um ano de diferença entre as notas é superior para o caso das escolas públicas. No primeiro quintil, os coeficientes passaram para 0,35 e 0,66, respectivamente, para as escolas públicas e privadas. No terceiro quintil, 0,64 e 0,82. No quinto quintil, 0,76 e 0,90.

Uma possível explicação para estes resultados é que as escolas particulares apresentam características mais estáveis de um ano para outro, como currículos mais bem definidos, corpo docente com menor rotatividade, metas de longo prazo e/ou práticas pedagógicas mais consolidadas (menos voláteis). Outra possível explicação é o fato de as escolas privadas poderem adotar uma política deliberada de escolha dos seus alunos. Dessa forma, o perfil do corpo discente é mais homogêneo ao longo do tempo.¹⁷

Entretanto, Kane e Staiger (2002) alertam para o fato de que, como não se efetua nessa abordagem nenhum controle para características dos alunos prestantes, é também possível que essa maior persistência associada a escolas particulares seja simplesmente o reflexo, ao longo do tempo, de uma maior estabilidade nas características da população que alimenta essas escolas, e não uma consequência intencional de práticas de ensino.

¹⁷ Agradecemos a um dos pareceristas que sugeriu essa explicação.



Fonte: Inep. Dados trabalhados pelos autores.

O próximo e último passo desta seção é decompor a variação da nota da escola no ENEM para quantificarmos a importância relativa de fatores persistentes e não persistentes. Confrontamos o comportamento de escolas públicas e privadas, além de separar a análise por diferentes quintis (por número de alunos), para comparar o efeito relativo desses componentes sobre escolas de diferentes portes.

A metodologia seguida aqui é a mesma de Kaine e Staiger (2001 e 2002), que sugerem que a variância total das notas de testes de proficiência (σ^2_y , ou “variância interescolar”), como o ENEM, pode ser decomposta em duas fontes: $\sigma^2_y = \sigma^2_{\text{persistente}} + \sigma^2_{\text{não persistente}}$, sendo $\sigma^2_{\text{persistente}}$ e $\sigma^2_{\text{não persistente}}$, respectivamente, os componentes de natureza persistente e não persistente. Adicionalmente, o componente de natureza não persistente pode ser sub-decomposto em dois: $\sigma^2_{\text{não persistente}} = \sigma^2_{\text{amostral}} + \sigma^2_{\text{resíduo}}$. Ou seja, parte do componente não persistente está diretamente relacionada ao fator amostral, que surge porque, dentro de uma mesma escola, o desempenho dos alunos raramente é uniforme.

Assim, a variância total da nota entre escolas é decomposta em três fatores – um componente persistente, um componente volátil e decorrente da variância amostral (intraescolar) e um componente volátil e decorrente de outros fatores transitórios. Ou seja, $\sigma^2_y = \sigma^2_{\text{persistente}} + \sigma^2_{\text{amostral}} + \sigma^2_{\text{resíduo}}$.

O processo de decomposição da variância, utilizando o resultado das provas objetivas do ENEM, foi aplicado separadamente para seis grupos de escolas do Estado de São Paulo, todas para o ano de 2006: primeiro, terceiro e quinto quintis das escolas públicas e das escolas privadas.

Os componentes da variância são calculados da seguinte forma:

A variância total, σ^2_y , é a variância das notas obtidas pelas escolas do seu grupo. A variância relacionada a fatores persistentes, $\sigma^2_{\text{persistente}}$, é calculada da seguinte forma: $\sigma^2_{\text{persistente}} = (\sigma^2_y) (\rho_1 / \hat{\rho}_{\text{persistente}})$, sendo ρ_1 a correlação entre as notas das escolas, dentro do mesmo grupo, um ano a frente (no caso, entre os anos de 2006 e 2007) e $\hat{\rho}_{\text{persistente}}$ a estimativa da correlação serial de primeira ordem do componente persistente.

Por sua vez, $\hat{\rho}_{persistente}$ é estimado como sendo a média simples dos termos (ρ_{k+1} / ρ_k) , para $k = 1$ a 3, para cada grupo de escolas. No caso, ρ_1 , ρ_2 e ρ_3 são a correlação entre as notas das escolas, dentro do mesmo grupo, respectivamente, um ano (entre os anos de 2006 e 2007), dois anos (entre os anos de 2006 e 2008) e três anos a frente (entre os anos de 2006 e 2009). Sendo assim, a fração $(\rho_1 / \hat{\rho}_{persistente})$ na expressão para $\sigma^2_{persistente}$ pode ser interpretada como uma estimativa do tamanho relativo do componente persistente, expressa como porcentagem, dentro da variância total medida no ano-base (2006).

A variância amostral, $\sigma^2_{amostral}$, foi gerada como a média simples entre as variâncias amostrais de cada escola do grupo. Essa etapa envolveu cálculos a partir de microdados (os desempenhos individuais de cada aluno): a variância amostral da escola “i” foi calculada como

$$\sigma^2_{amostral, i} = \frac{\sigma_i^2}{n_i}, \text{ onde } \sigma_i^2 \text{ é a variância intraescolar das notas dos}$$

“n_i” alunos da escola que realizaram a prova em 2006.¹⁸

Por fim, para cada grupo de escolas, o componente associado aos demais fatores não persistentes, $\sigma^2_{residuo}$, foi calculado como o resíduo na equação acima $(\sigma^2_y - \sigma^2_{persistente} - \sigma^2_{amostral})$.

Os resultados estão na Tabela 1. A Tabela 1(a) mostra as frações percentuais da variância das notas no ENEM que são explicadas por “fatores persistentes” e por “fatores não persistentes”, para o primeiro, terceiro e quinto quintis das escolas públicas; a Tabela 1(b) mostra as mesmas porcentagens para as escolas privadas. No caso de “fatores não persistentes”, ainda isolamos os efeitos decorrentes de pura “variação amostral” e de “outros fatores”.

Alguns resultados interessantes merecem destaque. Primeiro, em todos os grupos analisados, o componente “fatores persistentes” sempre tem a maior participação (acima de 50%) na decomposição da variância das notas, em relação a “fatores não persistentes” (a única exceção é o primeiro quintil das escolas públicas, com uma participação de “fatores não persistentes” de 60,56%). Como já apontado anteriormente, parte desta persistência está provavelmente associada a características estáveis do corpo discente das escolas, visto que não

¹⁸ Para detalhes sobre os dados utilizados no cálculo da variância amostral, ver seção anterior.

existe um ajuste para os níveis iniciais de desempenho dos alunos ou para as características demográficas dos mesmos (e não necessariamente a práticas educacionais persistentes nas escolas). No entanto, mesmo nas escolas de grande porte (5º. quintil), o componente transitório atinge patamares nada desprezíveis, de quase 15% para as escolas públicas, e 10% para as privadas.

Segundo, tanto para as escolas públicas como privadas, o peso dos “fatores persistentes” na decomposição da variância das notas do ENEM aumenta com o tamanho da escola. No caso das escolas públicas, o peso varia de 39,44% para o primeiro quintil até 84,07% para o quinto quintil. No caso das escolas privadas, esse peso varia de 62,57% para o primeiro quintil até 90,51% para o quinto quintil. O peso dos “fatores não persistentes” é maior para as escolas do primeiro quintil, pública (60,56%) e privada (37,43%), e no terceiro quintil das escolas públicas (33,01%). Em termos práticos, quando se comparam as notas atribuídas a duas escolas pertencentes a uma mesma dessas categorias, deve-se ter em mente que parte significativa da diferença que existe entre elas é transitória e pode desaparecer na medição seguinte.

Terceiro, com exceção das escolas de maior porte (quinto quintil), o peso da variação amostral é o mais importante dentro dos “fatores não persistentes”. Este ponto é interessante, pois a divulgação de listagens de médias por escola tende a destacar as diferenças de notas entre as instituições (variância *interescolar*). Essa abordagem usualmente tem como efeito colateral negligenciar as diferenças de notas *entre* alunos de uma mesma escola (variância *intraescolar*): a falsa impressão é que, dentro de cada instituição, o desempenho dos alunos é praticamente homogêneo. Apesar de esse ser realmente o caso nas escolas de grande porte, nas demais, a variância *intraescolar* de notas indica diferenças consideráveis de desempenhos entre um “bom” e um “mau” aluno dentro da mesma instituição. No caso extremo das escolas particulares de pequeno porte, 87,1% (= 32,61% / 37,43%) do componente não persistente da diferença de desempenhos entre instituições é decorrente de diferentes graus de desempenho entre alunos de uma mesma escola.

Por fim, em todos os quintis, o peso dos “fatores persistentes” é maior para as escolas privadas do que para públicas. No entanto, esta diferença cai com o tamanho da escola. Ele é 58,6%, 23% e 7,7% maior, respectivamente, no primeiro, terceiro e quinto quintis.

A análise da decomposição da variância das notas do ENEM sugere que é bastante questionável o uso dessas notas para comparar o desempenho das escolas. Esse alerta é agravado entre as escolas de menor porte e entre as públicas. No entanto, mesmo as escolas privadas de grande porte apresentam nada desprezíveis quase 10% da variância da nota do ENEM explicada por “fatores não persistentes”. Esses novos resultados corroboram os demais resultados desta seção quando analisamos gráficos de dispersão com as notas (ou ganhos) do ENEM contra o tamanho das escolas e a correlação serial entre as notas do ENEM de um até três anos à frente: é questionável a pertinência de utilizar as notas do ENEM para mensurar a diferença de desempenho entre as escolas e ao longo do tempo, principalmente para as escolas de menor porte. Esses resultados colocam em dúvida a utilização das informações contidas em um ranking baseado no ENEM por parte dos *stakeholders* das escolas nas suas tomadas de decisão.

Tabela 1: Decomposição da Variância das Notas do ENEM

Tabela 1(a)			
ESCOLAS PÚBLICAS:	1º. QUINTIL	3º. QUINTIL	5º. QUINTIL
FATORES PERSISTENTES	39,44%	66,99%	84,07%
FATORES NÃO-PERSISTENTES	60,56%	33,01%	15,93%
VARIAÇÃO AMOSTRAL	46,12%	21,43%	2,17%
OUTROS FATORES	14,44%	11,58%	13,75%

Tabela 1(b)			
ESCOLAS PARTICULARES:	1º. QUINTIL	3º. QUINTIL	5º. QUINTIL
FATORES PERSISTENTES	62,57%	82,43%	90,51%
FATORES NÃO-PERSISTENTES	37,43%	17,57%	9,49%
VARIAÇÃO AMOSTRAL	32,61%	16,61%	0,51%
OUTROS FATORES	4,82%	0,96%	8,98%

Fonte: Inep. Dados trabalhados pelos autores.

4. Previsibilidade e estabilidade do ranking do ENEM

O objetivo desta seção é testar se os *stakeholders* das escolas deveriam (ou não) levar em consideração a informação contida no ranking baseado nos resultados do ENEM em determinado ano, nas suas tomadas de decisões. Por exemplo, no momento em que escolhe uma escola para o seu filho, uma informação que interessa a um pai é saber se uma posição de destaque alcançada por essa escola no ranking, baseada no resultado dessa instituição no ENEM, é temporária ou se, por outro lado, tende a se manter no futuro, visto que o ensino médio tem duração de três anos (em termos técnicos, interessa ao pai ter alguma referência acerca da estabilidade intertemporal dessa medida). Ao mesmo tempo, é importante que ele saiba se o ranking hoje sumariza todas as informações relevantes disponíveis de forma eficiente ou se ele captura simplesmente efeitos transitórios e não duradouros.

Assim como os pais (e os alunos) estão interessados nestas respostas, pode-se dizer o mesmo de professores/diretores, seja para anteciparem possíveis consequências negativas ou positivas de novas colocações no ranking, seja para identificarem as escolas de maior prestígio para trabalhar. Essas informações são obviamente muito relevantes para outros *stakeholders*, como *policymakers* que porventura desejem identificar as características das melhores e das piores escolas, com fins de aprimorar a qualidade do ensino; ONGs que tenham interesse em focar suas atividades nas escolas com pior desempenho; investidores que queiram comprar escolas que estejam utilizando de forma ineficiente os recursos disponíveis; potenciais doadores que queiram “adotar” uma escola etc. Nesse contexto testamos, de duas maneiras, a qualidade desse ranking escolar.

O primeiro teste segue a metodologia utilizada em Dichev (1999), que investiga a previsibilidade de mudanças futuras nos rankings dos cursos de MBA dos Estados Unidos. O pressuposto da análise é o seguinte. Caso um determinado ranking sumarie todas as informações relevantes existentes de forma eficiente, então mudanças no ranking só podem ocorrer quando novas informações passam a ser disponibilizadas (que, por definição, não estão disponíveis quando o ranking foi criado). Desta forma, a incapacidade de prever mudanças futuras implica que o ranking corrente é uma boa aproximação de rankings futuros. Nas palavras de Dichev (1999), “finding

substantial predictability in changes allows one to conclude that current rankings are poor forecasts of future rankings”.

Como em Dichev (1999), é utilizada autocorrelação de primeira ordem nas mudanças das notas das escolas no teste do ENEM como uma *proxy* empírica da previsibilidade. Em outras palavras, avalia-se se as diferenças das notas médias por escola no ENEM num determinado período explicam essas mesmas diferenças em um período posterior. Caso isto ocorra, as mudanças no ranking do ENEM tem um componente transitório significativo que desaparece rapidamente na edição seguinte.

O segundo teste utiliza a metodologia de distribuição dinâmica, utilizada em outro contexto, por exemplo, em Laurini *et al.* (2005).¹⁹ A evolução da posição relativa de uma escola no ranking baseada nas notas do ENEM é modelada como um processo de Markov de primeira ordem. Por exemplo, calcula-se a probabilidade de uma escola estar, no futuro, entre as 10% melhores do país (ou em qualquer outra posição relativa no ranking) ou de um município, dado que ela estava neste patamar hoje. Caso esta probabilidade seja baixa, os *stakeholders*, ao realizarem suas escolhas, não deveriam dar um peso importante nas informações contidas no ranking hoje, pois ele não é capaz de antecipar a posição relativa da escola no futuro. Em outras palavras, o ranking provavelmente capturaria efeitos transitórios e não duradouros. Portanto, a qualidade das informações disponíveis no ranking seria de baixa qualidade, tendo em foco a escolha de uma escola, hoje, para um curso que tem duração prevista de três anos, como o ensino médio.

4.1. Previsibilidade das mudanças no ranking

Nesta subseção, testamos a qualidade do ranking do ENEM utilizando a correlação de primeira ordem nas mudanças das notas das escolas no teste do ENEM como uma *proxy* empírica da previsibilidade, como em Dichev (1999). Com este fim, foram tomadas as notas médias por escola na parte objetiva do ENEM para as edições de 2008, 2009 e 2010, tendo como universo as escolas do país. Com isto, obtivemos as seguintes séries:

¹⁹ Sobre essa metodologia, ver Quah (1996).

$$DIF_{1,i} = NOTA_{2009,i} - NOTA_{2008,i},$$

$$DIF_{2,i} = NOTA_{2010,i} - NOTA_{2009,i},$$

onde $NOTA_{t,i}$ é a nota média no ano “t” da escola “i” e $DIF_{1,i}$ e $DIF_{2,i}$ são as diferenças entre as notas da escola, respectivamente, para os anos 2009 e 2008 e para os anos 2010 e 2009.²⁰

A Tabela 2 apresenta a estatística descritiva. Ela mostra que, do período 2009/2008 para o período 2010/2009: (i) 17,5% das escolas obtiveram aumentos consecutivos nas suas notas médias; (ii) 32,9% das escolas obtiveram um aumento seguido por um decréscimo; (iii) 32% das escolas obtiveram um decréscimo acompanhado por um aumento; e (iv) 17,6% das escolas obtiveram decréscimos consecutivos.

Tabela 2: Manutenção de sinal X inversão de sinal em diferenças pareadas de primeira ordem

mudança POSITIVA seguida por mudança POSITIVA:	1.490 ocorrências (17,5%)
mudança POSITIVA seguida por mudança NEGATIVA:	2.798 ocorrências (32,9%)
mudança NEGATIVA seguida por mudança POSITIVA:	2.727 ocorrências (32,0%)
mudança NEGATIVA seguida por mudança NEGATIVA:	1.496 ocorrências (17,6%)

Fonte: Inep. Dados trabalhados pelos autores.

As situações de inversão de sinais representam 64,9% das ocorrências, contra 35,1% de manutenção de sinais. Essa evidência sugere que a maioria das mudanças (ganhos ou perdas) na nota padronizada de um ano para o outro são revertidas, total ou parcialmente, no período seguinte. Esse é um indício de “retorno à média”, o que, por extensão, sugere que as medições realizadas são continuamente contaminadas por ruído estatístico.

Existe outro forte indício de “retorno à média”. O coeficiente de correlação entre as mudanças nas notas das escolas ocorridas de 2010 em relação a 2009 e as mudanças ocorridas de 2009 em relação a 2008 é igual a -0,4667. Segundo Kane e Staiger (2002), “if a change in test scores is due to a one-time phenomenon, then the school is likely to revert back to its prior performance level in the next year. In the extreme case, if school test scores were pure noise and independent from one year to the next, we would expect a correlation

²⁰ As notas foram padronizadas em cada ano.

between the change this year and the change in the next to be $-0,5$ ". Este parece ser o caso aqui. A estimativa sugere que 93% da variância da mudança da nota da escola no ENEM é transitória ($-0,4667$ é 93% de $-0,5$). Dichev (1999) ainda aponta que o sinal negativo indica a tendência de retorno à média ("reversibility") nas séries temporais observadas para o universo de escolas, e que esse sinal registra a natureza oscilatória da medida empregada.

Se a hipótese de retorno à média for correta, também podemos esperar que as diferenças entre os scores de 2009 e 2008 carregassem algum poder explicativo sobre as diferenças entre os scores de 2010 e 2009, para cada instituição. Assim, efetuamos uma regressão entre as séries $DIF_{2,i}$ e $DIF_{1,i}$, utilizando Mínimos Quadrados Ordinários, na expectativa de medir a magnitude do efeito da mudança da nota num período sobre a mudança da nota no período seguinte. O modelo é um processo autorregressivo de primeira ordem na mudança das notas. Formalmente, temos $DIF_{2,i} = \alpha \cdot DIF_{1,i} + \varepsilon_i$, sendo " α " o coeficiente a ser estimado e ε_i o erro aleatório.

Obtivemos um coeficiente $\alpha = -0,469571$, estatisticamente diferente de zero (teste $t = -48,69089$), e um R^2 ajustado = $0,217889$. Mais uma vez, o sinal negativo indica que mudanças em um ano tendem a uma forte reversão no ano seguinte para o conjunto de escolas.

Os resultados desta seção sugerem que o ranking baseado no ENEM exibe uma autocorrelação, significativa e negativa, nas mudanças das notas. Isto implica que as mudanças no ranking têm um componente transitório importante, que desaparece rapidamente na edição seguinte do ranking.

4.2. Matriz markoviana

Nesta subseção, modela-se a evolução da posição relativa de uma escola no ranking baseado nas notas do ENEM como um processo de Markov de primeira ordem. O modelo utilizado é o seguinte:

$$\phi_{t+1} = P' \cdot \phi_t$$

onde ϕ e P são matrizes, respectivamente, de ordem $(n \times 1)$ e $(n \times n)$. Cada linha i da matriz ϕ_k ($k = t, t + 1$) mostra o número de escolas

no quantil i ($i = 1, \dots, n$) da distribuição das escolas no ranking no período k . Portanto, ϕ tem, por hipótese, um número finito de estados, sendo que cada estado corresponde a um determinado quantil da distribuição das escolas no ranking. Temos um processo de Markov de primeira ordem²¹ porque é assumido que a matriz ϕ no período $t + 1$ depende exclusivamente da densidade no período imediatamente anterior, t .

A matriz P é um operador que mapeia a transição entre a posição relativa de uma escola no ranking entre os períodos t e $t + 1$. O elemento P_{ij} da matriz indica a probabilidade de que o estado i seja seguido pelo estado j e, obviamente, $P_{i1} + P_{i2} + \dots + P_{in} = 1$. A estimativa para o elemento P_{ij} é igual a razão entre o número de escolas que estavam no quantil i e passaram para o quantil j no período $t + 1$ e o número de escolas que estavam no quantil i no período t . Obviamente, cada elemento da matriz é um número no intervalo $[0,1]$.

Os elementos da diagonal principal da matriz P (onde $i = j$) representam as proporções de instituições que permaneceram em seus respectivos quantis ao longo do tempo. Assim, o comportamento da diagonal principal é uma importante medida da estabilidade intertemporal do ranking. Quanto mais distante da diagonal principal se situa o elemento na matriz, mais radical é a migração de categoria. Isso pode ser interpretado como "mais instabilidade possui o ranking em questão" e, portanto, de menor qualidade ele é para antecipar a posição relativa de uma escola no futuro.

Antes de partir para os resultados, três explicações se fazem necessárias. Primeiro, a análise via matrizes de transição de Markov envolve necessariamente uma discussão sobre a qual o número de estados (a ordem da matriz P) ou sobre a segmentação mais adequada. Em nosso caso, optamos pela segmentação a 10% (decis), mas essa decisão, longe de ser um consenso ou padrão técnico, sempre incorpora alguma arbitrariedade e diferentes opções por tamanho de segmentação podem levar a conclusões diferentes. Para contornar essa limitação, refizemos a análise usando segmentações a 5% e a 20% (respectivamente, vintis e quintis) para testar a robustez dos resultados, sem alterações significativas nessas saídas.²²

²¹ A equação acima representa um modelo análogo ao de autorregressão de primeira ordem no qual os pontos são substituídos pelas distribuições completas. Para maiores detalhes, ver Laurini *et al.* (2005).

²² Os resultados com estas outras segmentações não foram colocadas nesta versão. Os autores

Segundo, com a escolha da segmentação por decis, é importante manter nas bases de dados apenas quantidades de escolas múltiplas de 10. Essa medida visa a garantir que os decis sempre contenham o mesmo número de escolas, evitando eventuais problemas de arredondamento decorrentes da discretização. Para tanto, as bases de dados foram truncadas até o múltiplo de 10 mais próximo, mas com perda mínima de observações. Optamos por eliminar, quando necessário, a(s) escola(s) de pior(es) desempenho(s) na avaliação mais recente da série temporal.²³

Por fim, assim como na seção anterior, as análises nesta seção utilizam, na sua quase totalidade, como referência para a construção dos rankings, as notas médias dos alunos de cada escola na parte objetiva da prova do ENEM. Uma comparação entre ambos os critérios será feita no final desta seção.

A Tabela 3 mostra a matriz de transição markoviana para todas as escolas brasileiras (total de 15.170 escolas) entre os anos de 2008 e 2009.²⁴ Alguns pontos podem ser destacados a partir do exame desta matriz.

Em cada linha da matriz, o maior valor é aquele que se encontra na diagonal principal. Isto significa que, não importa em qual decil a escola esteja, o mais provável é que ela se encontre no mesmo decil, em comparação com qualquer outro decil isoladamente.

Os números da diagonal principal giram entre 0,221 (decis 6 e 7) e 0,815 (decil 1). Por exemplo, em relação ao decil 1, esse último número significa que existe uma probabilidade de 81,5% de que “uma escola situada entre as melhores de 2008” continue sendo “uma escola situada entre as melhores do ano seguinte”.

A análise das diagonais principais revela um padrão interessante. Os únicos decis cujos valores da diagonal principal são superiores a 0,5 são os decis 1, 2 e 10. Para as escolas dos demais decis, o mais provável é que elas não permaneçam em 2009 no mesmo decil que estavam em 2008. Isto significa que as escolas localizadas nas pontas

podem disponibilizar os resultados para os interessados.

²³ Ajuste semelhante foi feito com as segmentações de 5% e 20%.

²⁴ As matrizes markovianas para o mesmo período, utilizando somente as escolas do estado de São Paulo ou do município de São Paulo, apresentam resultados similares. Os autores podem disponibilizar os resultados para os interessados.

das distribuições têm maior propensão a permanecerem em seus decis originais, ou sofrerem menores deslocamentos de categoria, para cima ou para baixo. Em outras palavras, escolas excepcionalmente boas, em termos relativos, tendem a permanecer no topo. Simetricamente, escolas excepcionalmente ruins, em termos relativos, tendem a permanecer na base da distribuição. O mesmo não se verifica para as escolas dos decis intermediários.

Tabela 3-Matriz de Transição, Todas Escolas (15.170), Brasil, 2008/2009

Prova Objetiva		2009									
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
2008	D1	0,815	0,177	0,007	0,001	0,000	0,000	0,000	0,000	0,000	0,000
	D2	0,182	0,606	0,162	0,027	0,012	0,005	0,003	0,003	0,000	0,001
	D3	0,003	0,157	0,392	0,241	0,108	0,051	0,029	0,011	0,007	0,001
	D4	0,001	0,037	0,227	0,285	0,207	0,123	0,065	0,032	0,020	0,005
	D5	0,000	0,012	0,115	0,217	0,237	0,185	0,117	0,074	0,028	0,015
	D6	0,000	0,005	0,051	0,113	0,194	0,221	0,191	0,134	0,069	0,021
	D7	0,000	0,005	0,026	0,065	0,129	0,198	0,221	0,189	0,115	0,052
	D8	0,000	0,000	0,015	0,034	0,072	0,125	0,181	0,223	0,225	0,125
	D9	0,000	0,001	0,005	0,013	0,034	0,067	0,136	0,210	0,285	0,250
	D10	0,000	0,000	0,000	0,005	0,008	0,024	0,057	0,123	0,252	0,531

Fonte: Inep. Dados trabalhados pelos autores.

É pouco provável que o conjunto de todas as escolas do Brasil componha o universo relevante para pais/alunos, no momento de escolher a instituição de ensino. Normalmente, a busca se concentra nas escolas do município onde eles moram e, para a maioria dos pais, as escolhas se restringem ou às escolas privadas, para aquelas famílias com renda suficiente para pagar as mensalidades, ou às escolas públicas. Adicionalmente, o interesse não se limita somente à evolução do desempenho da escola de um ano para outro, mas, pelo menos, durante os três anos de duração do ensino médio.

Tendo em vista esses pontos, a análise em seguida se concentra nas escolas do município de São Paulo, por ser aquele que contém o maior número de escolas do país, e mostra a evolução da posição relativa das escolas entre os anos de 2006 e 2009, separadamente para escolas públicas e privadas.²⁵

²⁵ O mercado relevante para pais/alunos provavelmente não compreende todas as escolas dentro de um município grande como São Paulo, onde as dificuldades de locomoção são significativas. No entanto, ele é menos abrangente que o mercado de todas as escolas do Brasil, além de possuir um número significativo de escolas que permite a construção de uma matriz de transição markoviana.

As Tabelas 4 e 5 mostram as matrizes de transição markoviana entre os anos de 2006 e 2009 para, respectivamente, as escolas públicas (total de 520 escolas) e privadas (250 escolas) localizadas no município de São Paulo. Alguns pontos podem ser destacados a partir do exame destas matrizes.²⁶

Somente em quatro linhas da matriz das escolas públicas (1, 2, 3 e 5), e em sete da matriz das escolas privadas (1, 2, 4, 5, 6, 9 e 10), o maior valor é aquele que se encontra na diagonal principal: não vale aqui o ponto de que, não importa em qual decil a escola esteja, o mais provável é que ela se encontre no mesmo decil, em comparação com cada outro decil isoladamente. Adicionalmente, o único elemento na diagonal principal, em ambas as matrizes, superior a 0,5 é o associado ao decil 1 (0,577 na pública e 0,680 na privada). Por conseguinte, somente para as melhores escolas públicas e privadas, o mais provável é que elas permaneçam em 2009 no mesmo decil que estavam em 2006.

A separação das escolas em públicas e privadas permite agora identificar uma diferença na dinâmica dos dois universos: todos os elementos da diagonal principal são maiores na segunda matriz (escolas particulares) que na primeira (escolas públicas). Ao mesmo tempo, nas posições triangulares da segunda matriz, os elementos são nulos, ou comparativamente menores.

O fato de o universo de escolas particulares ser menor (apenas 250 escolas, contra 520 instituições públicas) acarreta um menor tamanho dos decis, de modo que seria sensato esperar por uma maior transitoriedade – para uma escola pública migrar de um decil para o outro, seria necessário subir (ou cair) em média 26 posições (contra 12,5 posições no caso das escolas públicas). Entretanto, as matrizes acima revelaram exatamente o oposto: existe maior estabilidade na base de dados menor (escolas particulares).

Essa menor mobilidade entre decis no universo de escolas particulares pode ser explicada da seguinte forma. Podemos conjecturar que a pressão de mercado, decorrente de um ambiente concorrencial, gere incentivos diferentes para esse grupo de escolas. Uma posição

²⁶ Não existe uma diferença significativa entre as matrizes de transição para as escolas públicas e privadas do município de São Paulo entre os anos 2008 e 2009 e entre aquelas entre os anos de 2006 e 2009. Como consequência, as primeiras não foram reportadas neste trabalho. Os autores podem disponibilizar esses resultados para os interessados.

relativamente melhor no ranking do ENEM pode servir de gatilho para um círculo virtuoso, no sentido de que escolas privadas de boa performance seriam recompensadas pelo mercado já no curto prazo, com maior demanda e maior disponibilidade de recursos (mais capital e maior oferta de alunos e profissionais de qualidade). Esses fatores são fundamentais para que a escola possa reinvestir em práticas escolares que produzam efeitos positivos, e principalmente persistentes, sobre a performance escolar futura. Em outras palavras, escolas particulares com performance relativamente melhor teriam acesso crescente a mais recursos, essenciais para mantê-las em posições superiores no ranking. Em contraste, escolas particulares de performance relativamente inferior, com relativo menor acesso a esses recursos, tenderiam a manter-se nas posições inferiores no ranking, pois enfrentariam maior resistência em um eventual esforço de recuperação.

Essa interpretação é consistente com a situação verificada para escolas públicas, em que a ausência/debilidade de pressões dinâmicas e permanentes por alta performance poderia acarretar uma distribuição mais equitativa e menos criteriosa (nesse sentido) dos recursos disponíveis – o que teria como efeito final uma maior fluidez entre as categorias. Assim, a ausência de uma política de bonificações/sanções (por exemplo, na forma de escalonamento diferenciado de verbas) pelos órgãos oficiais sobre as escolas públicas relegaria o incentivo por bom desempenho aos próprios alunos, em detrimento de consequências efetivas sobre os gestores das escolas.

Tabela 4 - Matriz de Transição, Escolas Públicas (520), Município de São Paulo, 2006/2009

		Prova Objetiva									
		2009									
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1		0,577	0,212	0,058	0,077	0,019	0,019	0,000	0,019	0,000	0,019
D2		0,192	0,231	0,192	0,154	0,096	0,058	0,077	0,000	0,000	0,000
D3		0,077	0,154	0,327	0,154	0,096	0,019	0,058	0,058	0,019	0,038
D4		0,058	0,154	0,096	0,077	0,077	0,173	0,173	0,077	0,096	0,019
D5		0,058	0,000	0,096	0,192	0,231	0,135	0,096	0,077	0,077	0,038
D6		0,019	0,115	0,038	0,173	0,135	0,135	0,077	0,192	0,077	0,038
D7		0,019	0,038	0,038	0,019	0,115	0,154	0,154	0,154	0,115	0,192
D8		0,000	0,058	0,019	0,038	0,096	0,135	0,135	0,135	0,231	0,154
D9		0,000	0,000	0,115	0,096	0,038	0,115	0,077	0,154	0,135	0,269
2006	D10	0,000	0,038	0,019	0,019	0,096	0,058	0,154	0,135	0,250	0,231

Fonte: Inep. Dados trabalhados pelos autores.

Tabela 5 - Matriz de Transição, Escolas Privadas (250), Município de São Paulo, 2006/2009

Prova Objetiva

2009

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	0,680	0,200	0,120	0,000	0,000	0,000	0,000	0,000	0,000	0,000
D2	0,200	0,360	0,240	0,120	0,000	0,000	0,040	0,040	0,000	0,000
D3	0,080	0,280	0,240	0,160	0,160	0,000	0,000	0,000	0,080	0,000
D4	0,040	0,040	0,080	0,360	0,200	0,160	0,080	0,000	0,040	0,000
D5	0,000	0,080	0,240	0,120	0,240	0,200	0,040	0,040	0,040	0,000
D6	0,000	0,040	0,040	0,120	0,200	0,200	0,160	0,120	0,080	0,040
D7	0,000	0,000	0,040	0,000	0,160	0,200	0,160	0,200	0,080	0,160
D8	0,000	0,000	0,000	0,040	0,000	0,240	0,240	0,160	0,160	0,160
D9	0,000	0,000	0,000	0,040	0,040	0,000	0,160	0,200	0,320	0,240
D10	0,000	0,000	0,000	0,040	0,000	0,000	0,120	0,240	0,200	0,400

Fonte: Inep. Dados trabalhados pelos autores.

Nesta última parte da seção é refeita a Tabela 5, a matriz de transição para as escolas particulares do município de São Paulo entre os anos de 2006 e 2009. No entanto, a diferença é que agora não utilizamos como referência para a construção dos rankings as notas médias dos alunos de cada escola na parte objetiva da prova do ENEM. Ao invés disso, optamos por usar o critério oficial de ranqueamento adotado pelo MEC, que combina a média entre a prova objetiva e a redação, com pesos iguais (50%) para ambas as provas. Esse, por ser o critério oficial, é que produz aquele ranking largamente utilizado e difundido na mídia – razão pela qual é, possivelmente, o mais influente e impactante sobre os pais (ou responsáveis) na escolha da escola dos seus filhos.²⁷ Esta matriz encontra-se na Tabela 6.

As diferenças entre as matrizes das Tabelas 5 e 6 são significativas. Quase todos os números da diagonal principal são significativamente menores, na Tabela 6, quando o critério oficial do MEC é utilizado para ranquear as escolas, em relação aos resultados da Tabela 5. O único elemento da diagonal principal que aumenta é o referente ao último decil, que aumenta marginalmente, passando de 0,4 para 0,44. Adicionalmente, nas posições triangulares da matriz da Tabela 5, os elementos eram, em grande maioria, iguais a zero. Isto não mais ocorre quando o resultado das escolas na prova de redação também é considerado na análise.

²⁷ Ver, por exemplo, Veja São Paulo (2010).

Na matriz da Tabela 6, somente em quatro linhas da matriz, o maior valor é daquele elemento que se encontra na diagonal principal. Nenhum elemento da diagonal principal é superior a 0,5. Portanto, para qualquer escola, de qualquer decil, o mais provável é que ela não permaneça em 2009 no mesmo decil que estava em 2006.

O seguinte exemplo ilustra a mudança de persistência/volatilidade do ranking de desempenho das escolas quando o critério deixa de ser baseado somente no resultado da prova objetiva do ENEM, e passa a incorporar também o resultado da redação.

Suponha que um pai escolha uma escola que esteja entre as 10% melhores²⁸ no ranking do ENEM de 2006, quando se leva em conta somente a nota da prova objetiva, para que seu filho inicie o ensino médio em 2007. Três anos depois, ao final de 2009 (ano projetado para conclusão dessa escolaridade) haveria uma chance de 68% de essa instituição permanecer no topo, uma chance de 20% de ela cair para o 2º decil e uma chance de 12% dela cair para o 3º decil. Pode-se dizer que esta escola tem um desempenho “persistentemente alto”.

Entretanto, quando refazemos a análise usando a matriz da Tabela 6, considerando também a nota da redação, a mesma decisão teria efeitos bastante distintos: há somente 28% de chance de a escola permanecer no topo, 24% de chance de queda para o 2º decil, 20% de chance de queda para o 3º decil, e não desprezíveis 28% de chance da instituição posicionar-se, três anos depois, em algum estrato entre o 4º e o 9º decil. Segundo este critério, pode-se dizer que o conjunto de escolas mostra um desempenho bastante volátil.

Previsivelmente, a inclusão da nota de redação no critério final de ranqueamento deve provocar um aumento de volatilidade, na mesma medida em que esse tipo de prova incorpora um grau maior de subjetividade, dado pelos corretores. Surpreendente é o elevado grau de volatilidade. O ranking oficial do MEC, incorporando o resultado da redação, parece capturar efeitos transitórios e não duradouros. A qualidade das informações nele contida, em um determinado ano, parece de pouca valia para guiar a escolha futura dos pais/alunos.

²⁸ O mesmo tipo de análise pode ser feito para qualquer outro decil, com resultados similares.

Tabela 6 - Matriz de Transição, Escolas Privadas (250), Município de São Paulo, 2006/2009

Prova Objetiva + Redação		2009									
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
2006	D1	0,280	0,240	0,200	0,040	0,040	0,040	0,040	0,040	0,080	0,000
	D2	0,120	0,240	0,240	0,120	0,120	0,080	0,080	0,000	0,000	0,000
	D3	0,200	0,120	0,080	0,200	0,040	0,200	0,080	0,080	0,000	0,000
	D4	0,160	0,080	0,120	0,240	0,080	0,040	0,080	0,120	0,040	0,040
	D5	0,120	0,160	0,120	0,080	0,120	0,080	0,120	0,080	0,080	0,040
	D6	0,040	0,040	0,040	0,040	0,120	0,200	0,160	0,120	0,160	0,080
	D7	0,000	0,080	0,040	0,040	0,160	0,160	0,120	0,120	0,200	0,080
	D8	0,000	0,000	0,080	0,120	0,200	0,040	0,160	0,160	0,120	0,120
	D9	0,080	0,040	0,040	0,080	0,120	0,040	0,080	0,160	0,160	0,200
	D10	0,000	0,000	0,040	0,040	0,000	0,120	0,080	0,120	0,160	0,440

Fonte: Inep. Dados trabalhados pelos autores.

5. Conclusões

Os resultados encontrados neste trabalho questionam a qualidade do ranking das escolas do ensino médio baseado no ENEM. Em primeiro lugar, o tamanho da escola afeta o desempenho da mesma no ENEM. Escolares menores tendem a apresentar resultados mais extremos, altos ou baixos, além de flutuações de curto prazo mais significativas. Escolas menores também apresentam, na decomposição da variância da sua nota no ENEM, um maior peso relativo de “fatores não persistentes”.

Em segundo lugar, o ranking baseado no ENEM não parece capaz de sumarizar todas as informações relevantes existentes de forma eficiente, pois ele parece captar continuamente efeitos transitórios e não duradouros. Os rankings entre períodos apresentam uma autocorrelação, significativa e negativa nas mudanças das notas. Ou seja, escolas que têm um incremento (ou decréscimo) no desempenho relativo, em um determinado período, tendem a reverter essa variação no período seguinte. A probabilidade de uma escola estar, no futuro, num determinado decil do ranking, dado que ela estava neste decil hoje, é relativamente baixa.

O ranking baseado no ENEM é amplamente divulgado pela mídia e constitui um ranking oficial, pois é publicado pelo MEC, o que garante alta credibilidade junto aos *stakeholders* (pais, alunos, diretores

de escola, *policymakers* etc.) das escolas de ensino médio no país. Por esse motivo, crê-se que seus resultados sejam usados, por exemplo, por pais e alunos por ocasião da escolha de uma escola; da mesma forma, diretores parecem sofrer pressões da comunidade escolar, em caso de um resultado considerado insatisfatório.

No entanto, tendo em vista os resultados encontrados neste trabalho, a sua utilização pode conduzir a escolhas inapropriadas e, conseqüentemente, gerar alocação ineficiente de recursos, pois a construção do ranking é pautada por informações de qualidade duvidosa. Por exemplo, uma escola pequena pode atingir temporariamente uma posição relativamente boa no ranking, devido a fatores fora do controle de seu corpo administrativo e, como consequência, ser premiada no ano seguinte com uma elevada demanda pelos seus serviços. Como essa performance é efêmera e não sustentada, existiria uma elevada demanda pelos seus serviços, inexplicável pela verdadeira qualidade da educação oferecida.

Modificações no ranking das escolas do ensino médio divulgado pelo MEC parecem necessárias. Os resultados deste trabalho sugerem que o uso de notas em testes anuais pode não ser apropriado para comparar o desempenho entre escolas e ao longo do tempo, em particular quando se trata de escolas de pequeno porte, com poucos alunos. Existem algumas alternativas. A primeira, como sugerida por Kane e Staiger (2002), é a utilização de uma média das notas em diferentes anos. A segunda é ter um ranking baseado no valor adicionado, que leva em consideração qual é de fato a contribuição da escola para o aprendizado do aluno. A terceira seria divulgar um conjunto de informações, dentre outros a qualidade da infraestrutura, do corpo discente e docente, sem necessariamente agrupá-los, e deixar os responsáveis escolherem a escola de acordo com aquelas informações que julgarem mais relevantes. Por fim, dada a importância das habilidades não-cognitivas para o desempenho de um estudante, tanto acadêmico quanto profissional, informação a respeito da capacidade da escola para desenvolver essas habilidades é de fundamental importância.

Por fim, um comentário adicional se faz necessário. Uma particularidade do ENEM em relação a outros exames padronizados é o fato de ele ser facultativo e sua participação ter aumentado nos últimos anos, em particular no período analisado aqui. Acreditamos que um

ranking baseado no ENEM continuará sofrendo dos mesmos problemas apresentados nesse trabalho mesmo que a prova, porventura, se tornasse obrigatória. Isto porque escolas de menor porte continuariam igualmente vulneráveis às oscilações de curto prazo. Seria interessante, em pesquisa futura, testar a validade desse argumento, quando uma parcela próxima à totalidade dos alunos realizarem de fato a prova do ENEM.

Referências

- ANDRADE, E. Rankings em educação: tipos, problemas, informações e mudanças; *Estudos Econômicos*, Vol. 41.2, p. 323-343, 2011.
- ANDRADE, E.; MOITA, R.; SILVA, C. *A Escolha da Faculdade pelo Aluno: Estimação da Demanda e Precificação dos Atributos*. Mimeografado, 2009.
- CAMARGO, B; CAMELO, R.; FIRPO, S. E PONCZEK, V.; “Test Score Disclosure and School Performance”; Mimeografado.
- DICHEV, I. How Good Are Business School Rankings? *The Journal of Business*, Chicago, Vol. 72, No. 2, pp. 201–213, 1999.
- KANE, T. E STAIGER, D. The Promise and Pitfalls of Using Imprecise School Accountability Measures. *Journal of Economic Perspectives*, Vol. 16, No. 4, Fall, pp. 91–114, 2002.
- KANE, T. E STAIGER, D. *Improving School Accountability Measures*, NBER Working Paper # 8156, March, 2001.
- LAURINI, M., ANDRADE, E. E PEREIRA, P. Income Convergence Clubs for Brazilian Municipalities: A Non-parametric Analysis. *Applied Economics*, 37, 2099-2118, 2005.
- MONKS, J. e EHRENBERG, R. The Impact of U.S. News & World Report College Rankings on Admissions Outcomes and Pricing Policies at Selective Private Institutions, NBER, Working Paper # 7227, 1999.
- QUAH, D. “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics”, *The Economic Journal*, 106, 1045–55, 1996.
- SOUZA, P. “A Revolução Gerenciada”; Editora Pearson Prentice Hall, 2005.