

A EDUCAÇÃO PÚBLICA AVALIADA A PARTIR DAS BASES DE DADOS DO SIMAVE E ENEM

Sabrina Amélia de Lima e Silva – Universidade Federal de Minas Gerais (UFMG)

RESUMO

Para se conhecer a realidade educacional de um país, são necessários dados de oferta educacional, acesso ao sistema de ensino, modalidades de ensino, fluxo dos alunos ao longo da trajetória escolar e desempenho escolar. No Brasil, esses dados têm sido coletados por diferentes órgãos oficiais, em consonância com a atual Constituição Federal e a Lei das Diretrizes e Bases da Educação Nacional (LDB-1996), que estabelece como parte do dever do Estado, em relação à educação, o levantamento de informações estatísticas, cabendo à União a tarefa de coletar, analisar e disseminar informações sobre a educação. Este trabalho tem como objetivo fazer uma análise da educação a partir das bases de dados do SIMAVE e do ENEM. Procura-se relacionar as notas obtidas nessas provas com possíveis fatores de influência no desempenho escolar como, por exemplo, escolas de origem (municipal ou estadual), renda dos pais, ambiente de estudo, interesse dos professores, dentre outros.

PALAVRAS CHAVE

Educação Pública; ENEM; SIMAVE; Integração.

PUBLIC EDUCATION EVALUATED FROM SIMAVE AND ENEM DATABASES

ABSTRACT

In order to know the educational reality of a country, data of educational offerings, access to the education system, methods of teaching, flow of students throughout school life and school performance are necessary. In Brazil, these data have been collected by different agencies, in line with the Federal Constitution and the Law of Guidelines and Bases of National Education (LDB-1996). This law establish, as part of the duty of the State, with respect to education, the collection of statistical information, while the Union has the responsibility of collecting, analyzing and disseminating information on education. The purpose of this study is to analyze education from SIMAVE and ENEM databases. We intended to relate the scores obtained in these tests and possible influencing factors in school performance, for example, schools of origin (local or state), parental income, learning environment, teachers' interest, among others.

KEYWORDS

Public Education; ENEM; SIMAVE; Integration.

1. INTRODUÇÃO

A transição da adolescência para a vida adulta é um processo caracterizado por inúmeras interações e tensões existentes entre esses dois universos, sendo que a escola de ensino médio é um dos pontos que em especial marcam essa passagem para a fase adulta (BUCHMANN e DALTON, 2002). Com relação às expectativas dos jovens brasileiros quanto ao ensino médio esses citaram principalmente “preparar-me para o mercado de trabalho” e depois “preparar-me para o vestibular”, indicando a dicotomia existente com relação ao ensino médio (CORTI e SOUZA, 2009). Assim, esse nível de ensino pode ser o ponto final da educação formal obtido pelo jovem ou pode ser uma etapa de formação propedêutica para o ensino superior (BUCHMANN e DALTON, 2002).

O tema educação pública tem sido abordado frequentemente por grande número de estudiosos da área e ainda por outros que desejam iterar-se a respeito do assunto. Neste contexto, surgem questões referentes à eficiência dos sistemas de ensino e dos possíveis fatores que influem na qualidade educacional. Segundo Afonso e Aubyn (2005), um sistema educacional pode ser considerado eficiente, como qualquer outra atividade, se o acréscimo no produto se faz a expensas de um gasto com insumos constante ou se obtém um mesmo nível de produto com um gasto em insumo menor. Aplicando tal visão à educação devemos entender que em nosso contexto, as notas obtidas nos exames podem ser consideradas produto e os insumos seriam professores qualificados e estrutura escolar adequada por exemplo.

Através das bases de dados utilizadas objetiva-se verificar a possível relação de alguns determinantes no desempenho escolar. Hanushek, Gomes-Neto e Harbison (1996), evidenciam que em algumas regiões fatores como mobiliário para estudantes e professores, livros textos, recursos audiovisuais, guias para professores e materiais de escritório estão relacionados positivamente com o desempenho estudantil. Tendo como base a transição entre a adolescência e a fase adulta, e partindo do pressuposto que conhecer a escola de ensino médio e os diferentes fatores que influenciam o desempenho acadêmico dos estudantes são pontos decisivos para o desenvolvimento do sistema educacional brasileiro, procurou-se desenvolver neste texto análises empíricas a partir de base de dados do Sistema Mineiro de

Avaliação da Educação Pública (SIMAVE) e do Exame Nacional do Ensino Médio (ENEM), do ano 2007.

Para conhecer o sistema público mineiro de Educação, a Secretaria de Estado da Educação de Minas Gerais criou em 2000 o Sistema Mineiro de Avaliação da Educação Pública, o SIMAVE, responsável pelo desenvolvimento de programas de avaliação integrados. Os resultados das avaliações realizadas servem de base para responder às necessidades de planejamento e ação educacionais, servindo à realidade da sala de aula e influenciando a definição de políticas públicas para a educação no Estado. A criação do SIMAVE tem por objetivo desenvolver programas de avaliação integrados cujos resultados forneçam informações importantes para o planejamento de ações em todos os níveis do sistema de ensino. Este programa aponta as prioridades educacionais tanto para professores, especialistas e diretores quanto para os gestores do sistema, sendo fundamental na definição de ações para uma educação eficaz. Pela relevância de suas informações, o SIMAVE é um pilar do Projeto Estruturador do Governo de Minas Gerais. Escolas das redes Estaduais e Municipais de Minas Gerais participam do SIMAVE. Dessa forma, alunos das primeiras séries do Ensino Fundamental bem como os do 3º ano do Ensino Médio são avaliados. Três diferentes programas compõem o SIMAVE: o PROALFA, o PROEB e o PAAE. No presente trabalho, usou-se o componente PROEB do SIMAVE, restringindo-o a análises do 3º ano do Ensino Médio. Os resultados do PROEB focalizam a escola e esta mesma linha de raciocínio será aqui usada.

Em paralelo, utilizaremos também a Base de dados do ENEM visando captar aspectos que poderiam não ser notados apenas com o SIMAVE-PROEB.

Criado em 1998, o Exame Nacional do Ensino Médio (ENEM) tem o objetivo de avaliar o desempenho do estudante ao fim da escolaridade básica. Ele é utilizado como critério de seleção para os estudantes que pretendem concorrer a uma bolsa no Programa Universidade para Todos (ProUni). Além disso, cerca de 500 universidades já usam o resultado do exame como critério de seleção para o ingresso no ensino superior, seja complementando ou substituindo o vestibular. A divulgação das médias do ENEM tem se revelado como importante elemento de mobilização em favor da melhoria da qualidade do ensino, auxiliando professores, diretores e demais dirigentes educacionais na reflexão sobre deficiências e boas práticas no âmbito da escola, dado que é possível avaliar o desempenho obtido pelos alunos em cada área de conhecimento. Cabe ressaltar, no entanto, que a

utilização dos resultados do ENEM deve ser considerada com cautela, diante do caráter voluntário do exame, aspecto este que requer duas observações. A primeira refere-se ao fato de que, para algumas escolas, a amostra de estudantes que participaram do exame é demasiadamente pequena, o que pode tornar sua nota média pouco representativa do conjunto de estudantes da escola. A segunda observação é que, mesmo para as escolas com alta taxa de participação no ENEM, a amostra dos alunos de cada instituição pode não representar o desempenho médio que a escola obteria caso todos os alunos participassem. Em termos técnicos, pode haver um viés na seleção amostral. Por exemplo, se os alunos do Ensino Médio que pretendem cursar o nível superior forem os mais interessados em realizar o ENEM e estiverem mais bem representados pelos melhores alunos de cada escola, então haverá uma distorção, para cima, da média do ENEM observada por escola.

2. OBJETIVOS

Através deste trabalho, deseja-se comparar o Enem com o SIMAVE/PROEB. Os dois objetivos principais do projeto são: 1) Discutir o desempenho escolar das escolas públicas em Minas Gerais a partir, principalmente, das duas bases de dados citadas; 2) Comparar as duas bases, uma vez que existe um viés no desempenho escolar dos alunos que fizeram o ENEM, devido à seletividade positiva mencionada acima, em relação a outros estudos censitários.

3. MATERIAIS E MÉTODOS

Neste contexto recorreu-se às bases de dados do SIMAVE em 2007 e aos microdados do Enem no mesmo ano. O Sistema Mineiro de Avaliação da Educação Pública (SIMAVE) tem como missão avaliar para avançar e é aplicado aos alunos de escolas públicas da rede estadual de Minas Gerais (para maiores detalhes ver <https://www.educacao.mg.gov.br/projetos/>). Essa base de dados contém várias informações referentes aos alunos que realizaram a prova. Além das notas ou proficiências obtidas nas disciplinas de Língua Portuguesa e Matemática, que determinam o desempenho escolar do

estudante, há também várias características socioeconômicas e de percepção subjetiva do aluno a respeito de seu processo de escolarização apresentados. Na base utilizada há ainda informações como a escola, o município, a Secretaria Regional de Ensino (SRE), a rede e o turno no qual o aluno está inserido. Após alguns ajustes e adequação de variáveis, a base de dados do SIMAVE ficou com aproximadamente 90 variáveis explicativas. O ENEM, produzido pelo INEP, avalia as competências e habilidades obtidas no ensino médio (para maiores detalhes ver www.enem.inep.gov.br). Além disso, vários fatores são cobertos por essa base como questões socioeconômicas, escolaridade dos pais, cor, raça, gênero, etc., e também variáveis referentes à escola.

À medida que o número de variáveis a serem consideradas aumenta, há uma necessidade proporcional de maior conhecimento da estrutura e das inter-relações das variáveis (HAIR et al. 2006). Para tanto, a Análise Fatorial, técnica da [estatística](#) multivariada, foi utilizada neste trabalho. Ela foi destinada a representar um processo multivariado por meio da criação de novas variáveis, derivadas das variáveis originais e, geralmente, em menor número. Em termos gerais, a análise fatorial aborda o problema de analisar a estrutura das inter-relações (correlações) entre um grande número de variáveis definindo um conjunto de dimensões latentes comuns, chamadas fatores. Portanto, a análise fatorial é uma técnica de interdependência na qual todas as variáveis são simultaneamente consideradas, cada uma correlacionada com todas as outras, empregando ainda o conceito da variável estatística, a composição linear de variáveis.

Para aplicação da análise fatorial, em primeira mão, agrupou-se as variáveis dos bancos segundo fatores comuns, uma vez que muitas dessas variáveis podem indicar fenômenos parecidos e serem muito correlacionadas. Foram criados quatro grandes grupos: socioeconômico, individual, família e professor para o SIMAVE. Para o Enem utilizaram-se componentes como Ambiente humano da escola, Dedicção e ética dos professores, Nível socioeconômico, Escolaridade da família e outros.

Após aplicarmos a Análise fatorial em grupos de variáveis com suposta correlação, obtivemos novos grupos de variáveis através das componentes obtidas, como detalhado na próxima seção. Obtidas novas variáveis que agora representam todo o conjunto de dados, foram feitas regressões com mínimos quadrados ordinários (MQO), onde a variável dependente era o desempenho escolar.

4. RESULTADOS: REPRESENTACIONES SOCIALES DE DOCENTES Y ALUMNOS NATIVOS ACERCA DE LOS ALUMNOS MIGRANTES EXTERNOS.

Nesta seção os resultados serão apresentados em duas partes, a primeira conta com os resultados para o SIMAVE e a seguinte para o ENEM.

4.1 Resultados com o SIMAVE

Segue uma análise econométrica que utiliza a técnica dos MQO. Foram feitas regressões onde a variável dependente era o desempenho escolar. Os valores estão representados em porcentagem da prova. Os resultados são apresentados na tabela 1. No SIMAVE, para cada disciplina considerada (Língua Portuguesa e Matemática), além do todo, são separadas as melhores e as piores escolas, formando três grupos de estudo. A variável resposta, em todos eles, foi a nota do estudante nas provas de Língua Portuguesa e Matemática, analisadas em separado. Os valores nulos para cada uma das provas foram descartados nos modelos. Deve-se ressaltar que esses resultados não devem ser interpretados puramente como relação causal entre as variáveis e o desempenho escolar e sim como uma análise sobre os fatores associados a esse desempenho. Além disso, discutir qualidade da escola somente a partir de desempenho escolar apresenta uma série de limitações (Felício e Fernandes, 2005), uma vez que a qualidade de ensino apresenta diversas possibilidades distintas de análise.

Pela análise da tabela 1 e da tabela 2 para o SIMAVE, podemos retirar uma série de conclusões. Valores em negrito indicam coeficientes significativos. Quando se considerou todas as escolas do banco, todos os coeficientes foram significativos. Com todas as variáveis obtivemos 121255 observações em Língua Portuguesa e 106248 observações em Matemática. Para os demais grupos, as observações ficaram em torno de 1000.

Com relação ao sexo, foi criada uma dummy: 1 para mulheres, 0 para homens. Para a variável cor foram feitos dois grupos. Um com os alunos Brancos e Amarelos, representados por 1 e outros com os alunos Pardos, Negros e Indígenas, representado por 0. Os resultados

mostram que em língua portuguesa as mulheres se mostraram melhores, fato verificado pelo valor positivo para o coeficiente de sexo em língua portuguesa. Em contrapartida, em matemática o valor obtido é negativo, indicando que neste caso, homens são melhores em matemática. A variável idade está inversamente correlacionada ao desempenho. Assim, quanto maior for a idade do aluno, pior seu desempenho e, portanto, esta variável em ambos modelos deve apresentar valor negativo. Para todos os modelos a variável idade foi significativa, indicando que mesmo controlando pelos demais fatores, alunos mais jovens possuíam melhor desempenho. A dummy para cor, quando com coeficiente positivo, como no caso de todas as escolas inclusas indica que os pretos, pardos e indígenas foram piores em desempenho que os brancos e amarelos.

Para a variável **Itens Casa** que foi obtida pela análise fatorial agrupando as variáveis Quantidade de banheiros em sua casa, Quantidade de rádio em sua casa, Quantidade de geladeiras em sua casa, Quantidade de televisão em cores em sua casa, Quantidade de máquina de lavar roupa em sua casa, Quantidade de automóvel em sua casa e Quantidade de DVD em sua casa; verifica-se que só obteve-se significância nas piores escolas em língua portuguesa. Neste caso, alunos que possuem mais dos itens que compõe tal variável tiveram pior desempenho. Isso indica que outros fatores correlacionados com o nível socioeconômico parecem ser mais relevantes. No item **Básico**, que indica se a rua é calçada e se há água na torneira, houve significância para os modelos com todas as escolas. Captando efeitos de urbanidade e também de nível socioeconômico. O fato de o aluno possuir ou não **computador** em casa e com acesso à internet foi significativo nos modelos com todas as escolas. Alunos com computador e acesso a internet obtiveram melhores notas que os demais, indicando uma dimensão que capta efeitos socioeconômicos e também culturais. O item **repetência** foi significativo e positivo em todos os modelos. Ainda se destaca os coeficientes elevados nesta variável. Espera-se que alunos com matrícula regular, ou seja, que nunca repetiram o ano tenha melhor desempenho que os demais. Assim, esta variável deve estar positivamente correlacionada já que houve uma padronização onde o que é considerado melhor, assume maior valor. Isto se verifica pelos coeficientes obtidos. Note que isso ocorre mesmo com a inclusão da idade do indivíduo. O item **escolaridade**, que leva em consideração os anos de estudo dos pais, foi significativo para os modelos com todas as variáveis; alunos de pais com maior escolaridade têm melhor desempenho que os demais, indicando um ambiente domiciliar mais propício ao aprendizado. A variável **professor**, obtida pela análise fatorial

através do agrupamento de O (A) professor (a) mostra interesse no aprendizado de todos os alunos?, O (A) professor (a) dá oportunidade de os alunos expressarem suas opiniões?, O (A) professor (a) está disponível para esclarecer as dúvidas dos alunos?, O (A) professor (a) continua a explicar até que todos os alunos entendam a matéria?, O (A) professor (a) se esforça para ajudar os alunos?, O (A) professor (a) ajuda mais a uns do que a outros alunos? e O (A) professor dá as notas de maneira justa?, foi significativa para as piores escolas em português e para as melhores em matemática. O coeficiente positivo indica que ações do professor, nestes casos, ajudaram no desempenho do alunado. Além disso, esse valor obtido foi relativamente alto porque se espera que boa parte do desempenho dos alunos possa ser explicado pelo professor. Para organização, verifica-se que nas piores escolas em matemática, tal fator revela-se importante. O alto valor positivo indica que alunos inseridos em meios mais organizados possuem melhor desempenho. O fato de o professor faltar ou não as aulas foi significativo nas melhores escolas em português e nas piores em matemática. O poder explicativo das variáveis foi relativamente parecido para todos os modelos, como pode ser verificado pelo coeficiente de determinação, variando entre 0,093 e 0,162.

Tabela 1 – Regressões para o SIMAVE, 2007

VARIÁVEIS	LÍNGUA PORTUGUESA			MATEMÁTICA		
	TODAS	MELHORES	PIORES	TODAS	MELHORES	PIORES
Constante	61, 505	57, 802	34, 576	70,025	75, 577	45, 598
Sexo	1, 282	0, 845	3, 195	-3,063	-2, 975	-1, 789
Idade	-1,135	-1, 543	-0, 636	-1,535	-2, 192	-1, 218
Cor	0,847	-0, 470	0, 315	-1,494	-0, 767	-0, 350
Itens Casa	-2,005	0, 373	-7, 191	-2,081	0, 547	-1, 389
Básico	2,196	-0, 328	1, 520	1,956	-2, 054	1, 188
Computador	1,919	1, 000	2, 783	1,78	0, 115	-1, 351

Fonte: Elaborado pelo autor.

Tabela 2 – Regressões para o SIMAVE, 2007

VARIÁVEIS	LÍNGUA PORTUGUESA			MATEMÁTICA		
	TODAS	MELHORES	PIORES	TODAS	MELHORES	PIORES
Repetência	4,663	6, 411	4, 048	5,946	3, 958	4, 293
Escolaridade	2,786	1, 278	-0, 199	3,003	0, 437	-1, 594
Professor	2,491	0, 538	3, 983	3,841	3, 671	3, 703
Organização	3,896	-0, 619	0, 207	3,061	0, 222	4, 784
Correção	2,62	0, 170	4, 367	3,704	-1, 685	-2, 375
Falta	-0,969	2, 664	1, 793	-1,703	1, 517	4, 532

Fonte: Elaborado pelo autor.

4.2 Resultados com o ENEM

Para o ENEM, temos os resultados para a disciplina de Língua Portuguesa em todo o Brasil. Podemos verificar pontos similares aos do obtidos com o SIMAVE.

Tabela 3 – Regressões para o ENEM

VARIÁVEL	PORTUGUÊS	OBJETIVA	VARIÁVEL	PORTUGUÊS	OBJETIVA
Constante	46,2	72,9	Diretoria	0,24	0
Municipal	1,77	1,56	Sala de professores	0,61	0,54
Sexo	3,1	-2,8	Lab. de informática	0,06	0,21
Cor	1,01	0,63	Lab. de ciências	0,15	0,21
Idade	-0,74	-0,92	Outros laboratórios	1,9	0,93
Trabalha	-0,42	-0,39	Biblioteca	0,11	0,11
Escolarid. Mãe	0,65	0,31	Piscina	1,78	0,27
Renda familiar	1,36	0,82	Conexão à Internet	0,18	0
Turno noturno	-0,04	-0,03	Observações	698094	698094

Fonte: ENEM, 2005; Censo Escolar, 2005.

Para a variável sexo, com valor positivo, verifica-se que, assim como o SIMAVE, resultados mostram que em língua portuguesa as mulheres se mostraram melhores. Para a cor, novamente os brancos e amarelos apresentarem melhor desempenho que demais. A variável idade apresentou coeficiente igual a -0,74 indicando que alunos mais novos ou com menor distorção idade/série apresentaram melhor desempenho. Itens como escolaridade e renda familiar também seguiram os mesmos padrões verificados com o SIMAVE. Outros itens do ENEM, mesmo não sendo abordados diretamente da mesma forma no SIMAVE, se

mostraram bastante correlacionados. É o caso de sala de professores, existência de laboratórios, biblioteca, piscina e conexão à internet que podem ser comparados aos itens professor e computador descritos para a base anterior. Nesta base, o coeficiente de determinação foi de 0,173 e as observações próximas dos 700 mil. As interpretações acima foram retiradas a partir dos resultados da tabela 3 apresentada abaixo.

5. DISCUSSÃO

A diretriz que norteou as análises descritas até aqui foi a de buscar conhecimentos que melhorem ou permitem compreender fenômenos que ocorrem na educação mineira, especificamente nas escolas públicas, com intuito de possibilitar meios para modificar a realidade educativa e solucionar um problema nacional.

Através da técnica de Análise Fatorial foram criados e apresentados componentes que condensam informações acerca de variáveis correlacionadas em ambas as bases. Após tal medida, reuniu as variáveis grandes grupos. Cada um destes grupos apresenta variáveis com características em comum. Depois se fez uma padronização e junção de variáveis de cada grupo de forma a se formar doze novas variáveis que foram as componentes explicativas dos modelos de regressão.

Foram analisados quais eram os fatores que eram correlacionados com o desempenho escolar. Para o SIMAVE, notou-se que mais jovens, com computador de acesso à internet, que nunca repetiram o ano e inseridos num contexto mais organizado tiveram melhor desempenho. Os homens tendiam a apresentar melhores resultados para matemática, e as mulheres para português. O interesse, pontualidade e comprometimento do professor também se mostraram positivamente correlacionado com o desempenho do alunado qualquer que seja a matéria. Apesar da base do ENEM utilizada ser para o todo o Brasil, verificou-se similaridade aos resultados apresentados em Minas Gerais, objeto de estudo do SIMAVE.

Se compararmos os resultados obtidos para o SIMAVE e ENEM notou-se que esse estudo aponta para algumas positivas correlações entre variáveis referentes à família, professor, socioeconômico e individual do aluno. A infraestrutura escolar avaliada no ENEM pela existência de laboratórios, biblioteca, área de lazer, sala de professores e diretoria também se relacionou com o SIMAVE indicando positiva correlação desses itens com o desempenho do alunado.

A partir deste ponto pode-se ampliar para estudos mais detalhados que norteiem a adoção de meios para modificar a realidade educativa de Minas Gerais e até mesmo de outros estados.

REFERENCIAS

BUCHMANN, C.; DALTON, B. **Interpersonal influence and educational aspirations in 12 countries: the importance of institutional context**, *Sociology of educations*, 2002. 75 p.

CORTI, A.; SOUZA, R.; **Que Ensino Médio queremos?** São Paulo: Ação Educativa, 2009.

AFONSO, A. E AUBYN, M.. **“Cross-Country Efficiency of Secondary Education Provision: A semi-parametric Analysis with non-discretionary Inputs”** European Central Bank Working Paper Series No. 494/June 2005. <http://www.ecb.int/pub/pdf/scpwps/ecbwp494.pdf> acesso 27/02/2006.

FELÍCIO, F.; FERNANDES, R.; **O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de São Paulo**, Em: Anais do XXXIII Encontro Nacional de Economia, ANPEC, 2005.

HANUSHEK, E. A, GOMES-NETO, J. B., HARBISON, R. W. Efficiency-enhancing investments in school quality. In: BIRDSALL, N., SABOT, R. H (Eds.) **Opportunity forgone: education in Brazil**. Washington, DC.: Inter-American development Bank. p.385-424,1996.

HECKMAN, J. **Sample selection bias as a specification error**, *Econometrics*, 1979. 47 p.

SABRINA AMÉLIA DE LIMA E SILVA

Acadêmica em Ciências Atuariais – Universidade Federal de Minas Gerais

Orientador: André Braz Golgher

E-mail: silva.saamelia@gmail.com