

Identificando Características Latentes de Escolas Públicas de Ensino Médio Mineiras

Identifying Public Secondary Schools Latent Characteristics in Minas Gerais

Geanderson Eduardo Ambrósio^{a*}; Talita Priscila Pinto^a; João Eustáquio de Lima^a

^aUniversidade Federal de Viçosa, Programa de Pós-Graduação Stricto Sensu em Economia Aplicada.

*E-mail: geanderson.ambrosio@ufv.br

Resumo

As características intrínsecas às escolas se correlacionam diretamente com a qualidade do capital humano por elas formado, o que torna o estudo destas propriedades parte fundamental no processo de universalização do ensino básico no Brasil. Na presente pesquisa, quatorze variáveis de 2177 escolas de ensino médio de Minas Gerais foram submetidas à técnica de Análise Fatorial Exploratória, o que permitiu a extração de quatro fatores (variáveis latentes) que as caracterizaram. O primeiro fator agregou características de dimensão física, o segundo e quarto fatores agregaram características infraestruturais, enquanto o terceiro fator agrupou variáveis relativas à qualificação do corpo docente e discente. A composição fatorial, analisada segundo localização da escola em espaço rural ou urbano, revelou que escolas rurais estão muito aquém quanto à infraestrutura básica, como esgotamento, coleta de lixo e fornecimento de água por rede pública. Por outro lado, observa-se menores discrepâncias quanto às autarquias municipal e estadual. Escolas municipais são maiores e têm melhor capacitação do capital humano, enquanto as estaduais são mais qualificadas estruturalmente.

Palavras-chave: Educação. Ensino Médio. Minas Gerais. Análise Fatorial.

Abstract

The intrinsic characteristics to schools are directly correlated to the human capital quality formed by them, which makes important the study of these properties in the universalization of elementary school process in Brazil. In this study, fourteen variables from 2177 secondary schools in Minas Gerais were submitted to exploratory factor analysis technique, which allowed the extraction of four factors (latent variables) that characterized them. The first factor grouped physical dimension features, the second and the fourth factors grouped infrastructural characteristics, while the third factor grouped variables relating to the faculty and students qualification. The factorial composition, analyzed according to the school location in rural or urban areas, revealed that rural schools fall far short as regard the basic infrastructure, such as sewage, garbage collection and public water supply. On the other hand, there are minor discrepancies in the municipal and state levels. Municipal schools are larger and have better human capital training, while state schools are structurally more qualified.

Keywords: Education. Secondary School. Minas Gerais. Factor Analysis.

1 Introdução

O modelo de ensino médio utilizado no Brasil tem sido modificado de tempos em tempos. A partir da década de 30, desenvolveram-se duas categorias de ensino: uma propedêutica, na qual se vislumbra o ingresso na universidade; e outra de natureza técnica, já articulada com o mercado de trabalho. De acordo com a Constituição Federal de 1988, em seu Artigo 205, a responsabilidade política pela educação recai sobre o estado e, face o não cumprimento deste dever, compete responsabilização criminal dos governos (CARNEIRO, 2010). Tal compromisso foi ratificado em 2009, com a publicação da Emenda Constitucional nº 59, que incumbiu obrigatoriedade e gratuidade da educação básica dos 4 aos 17 anos de idade a partir do ano de 2016, inclusive aos indivíduos que a ela não tiveram acesso em tal faixa etária (LIMA, 2011).

Dados do ENEM (Exame Nacional do Ensino Médio) 2012 indicam certa segregação no resultado do desempenho dos alunos. Escolas estaduais e municipais ocupam colocações inferiores em relação às federais e particulares. Além disso,

observa-se discrepância de desempenho favorável a alunos de escolas da zona urbana em detrimento da zona rural. A despeito da majoritariedade de matrículas no sistema público de ensino, não se observam políticas capazes de garantir o acesso ao ciclo completo da educação básica de qualidade que melhorem o desempenho daqueles menos favorecidos (COSTA, 2013).

Diversas realidades estruturais e acadêmicas estão relacionadas às características socioculturais dos alunos. Espera-se que o governo atue levando ensino de qualidade à totalidade dos alunos em idade escolar, mas não é isto que se observa. Nos anos dois mil houve estagnação e queda no número de matrículas, tendência contrária àquela observada na década de noventa (SIMÕES, 2009; KUENZER, 2010; COSTA, 2013).

A partir de 2007, foram desenvolvidas normas e medidas focadas na qualidade, como a constituição dos Institutos Federais – IF e do Programa Ensino Médio Inovador, além da reestruturação do ENEM em 2009 e 2010, que começou a ser utilizado como vestibular para universidades federais e como forma de induzir a organização curricular do ensino

médio (MELO; DUARTE, 2011; INEP, 2011a). Ainda assim, na literatura recente, não foram verificados sinais de que estes esforços tenham surtido efeitos claros sobre a qualidade ou o número de matrículas do ensino médio no Brasil.

Em 2004, a taxa de frequência líquida escolar foi de 93,8% para o Ensino Fundamental e 44,4% para o Ensino Médio (INEP, 2006) e, além disso, somente um em cada dois ingressantes no ensino médio concluem esta etapa da educação fundamental (SIMÕES, 2009), o que pode ilustrar um possível “gap” na relação de atratividade entre a criança e o jovem com a escola. Quanto às zonas rural e urbana, observa-se que a taxa de frequência líquida escolar para o ensino médio foi 22,1% na zona rural e 49,4% na zona urbana (INEP, 2007). Em 2008, pouco mais da metade dos jovens de 15 a 17 anos estavam matriculados nas redes de ensino médio e aproximadamente 15% nem mesmo estavam matriculados no sistema de ensino (LIMA, 2011). Essa baixa atratividade pode ser reflexo de escolas com provimentos precários de infraestrutura, professores pouco capacitados e, no caso da zona rural, condições socioeconômicas desfavoráveis, que acabam determinando resultados inferiores em termos de progressão escolar (INEP, 2006).

A Teoria do Capital Humano pressupõe que todos os cidadãos podem tomar decisões racionais e livres, há os que vão investir mais em sua própria educação e com isso vão conseguir maiores rendimentos, e aqueles que irão se acomodar em níveis de rendimentos inferiores. Alunos frequentadores de escolas que não fornecem o mínimo exigido em estrutura e corpo docente estão sujeitos a uma formação deficiente tanto para o mercado de trabalho quanto para o meio acadêmico. O sistema educacional em si irá responder apenas às demandas individuais, sem promover oportunidades mais igualitárias (CATTANI, 1997), de modo que a decisão quanto ao *trade off* entre estudo e trabalho é de responsabilidade do indivíduo, mas cabe ao governo minimizar o custo de oportunidade desta decisão.

São várias as características que podem descrever a realidade do meio escolar e auxiliar no entendimento das realidades com que discentes e docentes se deparam. O Censo Escolar, por exemplo, uma pesquisa de alcance nacional para os ensinos fundamental e médio, discorre sobre centenas de variáveis para alunos, professores, turmas e escolas. A compreensão dessas características mostra-se importante, já que é razoável presumir que elas estejam associadas à qualidade do capital humano formado na instituição.

No presente estudo as escolas mineiras municipais e estaduais de nível médio serão investigadas quanto a algumas destas variáveis. Espera-se contribuir para a caracterização geral da estrutura, corpo docente e discente, preenchendo assim parte da demanda por estudos que melhorem o entendimento da realidade escolar do ensino médio apontada por Kuenzer (2010). Para tanto, será empregada a técnica de Análise Fatorial Multivariada (HAIR *et al.*, 2009), na qual as variáveis (características) serão associadas de tal forma que

resultem em dimensões latentes (fatores) englobantes das variáveis estatisticamente semelhantes. Serão consideradas variáveis relativas ao corpo docente, discente e infraestrutura das escolas, e espera-se que estas variáveis agrupem-se em fatores semelhantes a estas dimensões. Os resultados da pesquisa permitirão inferir estatisticamente sobre a viabilidade de assumir tais dimensões como atributos gerais para caracterizar as escolas. Dessa forma, a análise estatístico-descritiva poderá contribuir com a discussão sobre o assunto e melhorar a compreensão das especificidades do meio acadêmico do ensino médio em Minas Gerais.

2 Material e Métodos

O presente estudo é uma abordagem empírico-analítica que envolve coleta, tratamento e análise de dados das escolas municipais e estaduais de ensino médio de Minas Gerais, a partir dos microdados do Censo Escolar 2012 (INEP, 2011b) e do ENEM 2012 (INEP, 2011b).

O universo de escolas de ensino médio mineiras engloba 2500 unidades, das quais 83% (2237 escolas) são estaduais ou municipais, e registram 91% do total de matrículas. Foram constatadas algumas observações cuja lógica dos dados era questionável ou as informações eram incompletas, as quais foram retiradas da amostra. Restaram então 2177 escolas para a análise, das quais 60 são de autarquia municipal e 2117 estadual; 130 estão localizadas em zona rural e 2047 em zona urbana. Os dados foram extraídos e processados pelos softwares IBM SPSS versão 16.0 e STATA versão 12.0 e serão apresentados pelas perspectivas de autarquia municipal/estadual e de localização rural/urbana. Para entender como as variáveis se relacionam estatisticamente, será conduzida a técnica de Análise Fatorial Exploratória segundo os procedimentos metodológicos de Hair *et al.* (2009).

Inicialmente, o teste de diferença entre médias de Hotelling testará a hipótese nula de que os vetores de médias das variáveis pelas perspectivas de autarquia e localização sejam estatisticamente iguais, a fim de verificar se as especificidades das escolas nestas dimensões são estatisticamente relevantes. Tal teste também será aplicado ao vetor de médias dos escores fatoriais, que serão extraídos pela técnica de Análise Fatorial. Será conduzida em sequência uma breve análise descritiva dos dados. Para testar a viabilidade desses dados para a Análise Fatorial, serão aplicados três procedimentos (HAIR *et al.*, 2009): uma matriz anti-imagem, que fornece o negativo das correlações entre as variáveis, as quais não devem ser superiores a 0,7; o Teste de Esfericidade de Bartlett, que testa a presença de correlações entre as variáveis com a hipótese nula de que a matriz de correlações seja uma identidade (ausência de correlações); e, finalmente, o Teste Kaiser-Meyer-Olking (KMO), uma medida de adequação dos dados ao método e que deve apresentar resultado maior que 0,5 para cada variável e para o ajuste geral do modelo (*overall*).

A Análise Fatorial - AF é uma técnica de interdependência

entre variáveis que tem o propósito principal de definir uma estrutura de dimensões latentes, denominadas fatores. Um número de fatores menor que o número de variáveis deve ser capaz de refletir as características do objeto sem perda relevante de informações (HAIR *et al.*, 2009; MINGOTI, 2005). Existem duas categorias distintas de AF, a Análise Fatorial Confirmatória - AFC e a Análise Fatorial Exploratória - AFE¹. A condução da presente pesquisa se dá pela AFE, pois não se verificaram estudos anteriores com aplicações da análise fatorial em perspectiva semelhante a desta pesquisa que pudessem embasar uma possível AFC.

A extração dos fatores se dá pela técnica da Análise de Componente, sendo considerados o critério *scree*, o critério de raiz latente e o princípio da interpretabilidade e da parcimônia. O critério *scree* pressupõe que grande parcela da variância será explicada pelos primeiros fatores e que entre os fatores haverá uma diferença numérica significativa; quando essa diferença torna-se baixa, esse ponto determina o número de fatores a ser extraído (MÁRIO, 2007). Graficamente, num plano de autovalores versus o número de fatores, esse é o ponto em que após uma inflexão a curva se aproxima de uma reta horizontal. O critério de raiz latente indica que devem ser considerados apenas os fatores cujos autovalores forem maiores que um (HAIR *et al.*, 2009). O princípio da interpretabilidade institui que os fatores extraídos devem ter um sentido dentro do contexto estudado, enquanto o princípio da parcimônia infere que o mínimo de fatores possível deve ser extraído (MINGOTI, 2005).

Serão analisadas também as comunalidades de cada variável, que representam a proporção da variância da variável que é explicada pelo conjunto dos fatores extraídos. Variáveis com comunalidade inferior a 0,5 podem ser retiradas da amostra, a fim de verificar se existe melhora nos outros indicadores de ajustamento dos dados.

A técnica de Análise de Componente considera a variância total dos dados e derivam-se fatores que contém pequenas proporções de variância única e, em alguns casos, variância de erro. O primeiro fator extraído é aquele que explica o máximo de variabilidade dos dados, o segundo fator explica menor variabilidade que o primeiro, e assim sucessivamente. A fim de simplificar a interpretação dos fatores estimados, utiliza-se a rotação ortogonal Varimax, que minimiza a ocorrência de uma variável possuir altas cargas fatoriais para diferentes fatores (MÁRIO, 2007). Os escores fatoriais serão estimados pelo método de Bartlett, que, apesar de produzir escores menos precisos do que quando estimados pelos Mínimos Quadrados Ordinários, possui a vantagem de que estes são não viesados (STATACORP, 2013).

As variáveis utilizadas se referem às características da estrutura da escola, dos docentes e dos alunos, e foram obtidas junto ao Censo Escolar 2012 e ao ENEM 2012. O Censo Escolar é um levantamento de dados estatístico-educacionais

de abrangência nacional, realizado anualmente, e é o principal instrumento de coleta de informações da educação básica (INEP, 2011c). O ENEM também é de abrangência nacional e é aplicado anualmente com o objetivo de avaliar o desempenho do estudante ao fim da educação básica (INEP, 2011a), contudo nem todos os alunos concluintes o realizam.

A utilização de resultados do ENEM como *proxy* para a qualidade do ensino pode ser controversa. Andrade e Soida (2013), por exemplo, questionam a confiabilidade do ranking das escolas do ensino médio baseado no ENEM, mas ainda assim reconhecem que o ranking do ENEM é o único indicador de desempenho, de abrangência nacional, para as escolas de ensino médio. Deste modo, na ausência de um indicador de melhor qualidade, o uso do resultado do ENEM é justificável e viabiliza a condução da pesquisa. O Quadro 1 apresenta e descreve as variáveis utilizadas na Análise Fatorial.

Quadro 1: Descrição das variáveis

| Variável | Descrição | Categoria |
|--------------|--|----------------------------|
| n_alunos | Número de matrículas | Numérica |
| n_func | Número de funcionários | Numérica |
| n_salas_util | Número de salas utilizadas | Numérica |
| superior | Percentual de prof. com curso superior | Numérica |
| esp_mes_doc | Percentual de prof. com pós graduação | Numérica |
| bbt | Biblioteca | 1 = presença; 0 = ausência |
| lab | Laboratório de ciências ou de informática | 1 = presença; 0 = ausência |
| net | Acesso à internet | 1 = presença; 0 = ausência |
| esgoto_rp | Coleta de esgoto por rede pública | 1 = presença; 0 = ausência |
| lixo_cp | Coleta periódica de lixo | 1 = presença; 0 = ausência |
| agua_rp | Abastecimento de água por rede pública | 1 = presença; 0 = ausência |
| raca | Percentual de alunos brancos e amarelos | Numérica |
| resid | Percentual de alunos residentes em zona urbana | Numérica |
| nota_obj_red | Média simples da nota dos alunos no ENEM | Numérica |

Fonte: Dados da pesquisa.

3 Resultados e Discussão

3.1 Resultados descritivos segundo localização e autarquia

O Teste de Hotelling de diferenças entre médias apontou que o vetor de médias das variáveis é diferente ao nível de significância de 1% para a segregação amostral por localização. As escolas do meio rural têm, em média, 100 alunos, enquanto as do meio urbano, 410; mas a proporção de funcionários por aluno é inversa, as escolas de meio rural apresentam 0,47 funcionários/alunos, enquanto as do meio urbano, 0,24. O

1 Para detalhes sobre as categorias de Análise Fatorial, ver Hair *et al.* (2009).

número médio de salas utilizadas é menor nas escolas da zona rural em relação às da zona urbana, respectivamente 7,7 e 12,6. O número médio de alunos por sala utilizada é de 12,9 na zona rural, enquanto é de 32,6 na zona urbana. Estes dados apontam para um sistema educacional de zona rural que apresenta menor ganho de escala que na zona urbana, pois são escolas com um número menor de salas e alunos, mas que possuem proporcionalmente mais funcionários. Considerando, por exemplo, a contratação de um professor por sala, o custo do professor por aluno nas escolas da zona rural tende a ser maior.

Das escolas do meio rural, em 66,92% delas todos os professores possuem curso superior completo, enquanto nas do meio urbano este número é de 72,59%. Quanto à pós-graduação, para as escolas do meio rural, em 16,15% delas nenhum professor possui curso de tal natureza, enquanto que para as escolas do meio urbano, em apenas 2,93% das escolas isto ocorre. Estes dados apontam para melhor capacitação dos professores que atuam nas cidades em detrimento do campo. Tal característica pode estar diretamente ligada ao melhor desempenho no ENEM dos alunos de escolas urbanas, comparando-os aos alunos de escolas rurais, que obtiveram nota média de 485,18 e 462,74, respectivamente.

Quanto à estrutura física das escolas, observa-se que novamente as escolas do campo são as mais deficientes: 74,62% delas não possuem esgotamento por rede pública; 47,69% não possuem coleta periódica de lixo e 60,77% não dispõem de fornecimento de água por rede pública.

Na contramão destes dados, destaca-se que todas as escolas da amostra são providas de energia elétrica pela rede pública, e tal situação inviabilizou a inclusão de uma possível variável referente à energia na AF devido à ausência de variância dos dados. Outro resultado satisfatório diz respeito ao acesso à internet, já que 99,36% das escolas urbanas e 90% das rurais o possuem. Ainda que os dados não especifiquem o tipo de acesso, esse resultado pode ser entendido como um processo de universalização da internet.

Quanto à autarquia administrativa das escolas, o Teste de Hotelling de diferenças entre médias também apontou que o vetor de médias das variáveis é diferente ao nível de significância de 1%, contudo nota-se menor discrepância dos dados do que nas dimensões de localidade. O número médio de alunos na rede municipal, por exemplo, é de 236, enquanto na rede estadual este número é de 396, mas a quantidade de funcionários por aluno na rede municipal é maior, 0,61, enquanto na rede estadual são 0,24. O número médio de salas utilizadas na rede municipal é de 15,2, enquanto na rede estadual é de 12,2 salas, o que leva a um número médio de 15,5 alunos por sala na rede municipal e 26 na rede estadual. Tais dados podem inferir que as escolas estaduais sobrecarreguem mais seus funcionários que as municipais, o que também pode refletir negativamente no desempenho dos alunos.

Quanto à formação docente, em 80% das escolas municipais todos os professores possuem curso superior, enquanto que na rede estadual esse número é de 72,04%. Em relação à pós-graduação, 8,33% das escolas municipais não apresenta nenhum professor com tal titulação, enquanto que na rede estadual esse

número compreende 3,59%. A nota das escolas no ENEM, discriminada para dependência administrativa, demonstra um resultado ligeiramente superior para as escolas municipais, 497,34 frente a 485,04 das escolas estaduais, uma diferença que pode ser relacionada à sobrecarga dos funcionários da rede estadual e à pior qualificação de seus professores.

Referindo-se à estrutura física das escolas, novamente, o esgotamento sanitário é o dado de maior relevância; foi observado que 11,67% das escolas municipais e 15,97 das estaduais não possuem esgotamento por rede pública. Contudo, essa característica pode estar mais relacionada à infraestrutura dos municípios, por não levarem rede de esgoto às suas zonas rurais, que propriamente à autarquia administrativa da escola. De um modo ou de outro, tal resultado aponta que o saneamento básico é um problema persistente nas escolas mineiras independentemente da ótica avaliada nesta pesquisa.

3.2 Análise de viabilidade dos dados e extração dos fatores

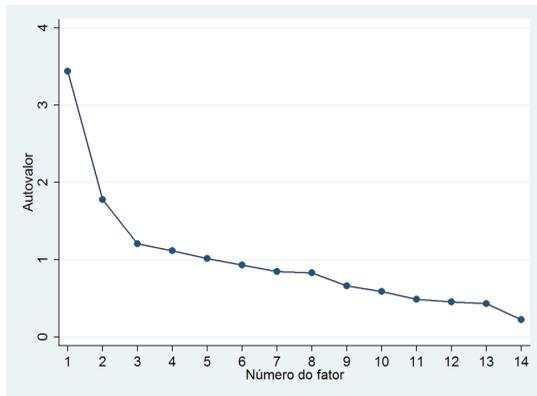
A matriz anti-imagem das variáveis não apresentou correlações parciais superiores a 0,7, o que é um indicativo de que os dados são adequados ao modelo. O Teste Esfericidade de Bartlett rejeitou a hipótese nula da matriz de correlações ser uma matriz identidade a 1% de significância, indicando a presença de correlações suficientes entre as variáveis para a aplicação da AF. O critério KMO apresentou valores superiores a 0,5 para cada uma das variáveis e um ajuste geral de 0,7805, o que também credencia os dados para a AF.

Para a determinação do número de fatores a serem extraídos, foram considerados o critério de raiz latente, o critério *scree* e a interpretabilidade e parcimônia. O critério de raiz latente indica a extração de cinco fatores, os quais possuem autovalor maior que um e explicam 61,08% da variância dos dados (Quadro 2). O critério *scree* indica a extração de três fatores, os quais explicam 45,86% da variância dos dados (Figura 1).

Quadro 2: Resultado do modelo de AFE

| Fator | Autovalor | Diferença | Proporção da variância explicada | Proporção acumulada |
|-------|-----------|-----------|----------------------------------|---------------------|
| 1° | 3,43702 | 1,65864 | 0,2455 | 0,2455 |
| 2° | 1,77838 | 0,5727 | 0,127 | 0,3725 |
| 3° | 1,20568 | 0,0913 | 0,0861 | 0,4586 |
| 4° | 1,11438 | 0,09909 | 0,0796 | 0,5382 |
| 5° | 1,01529 | 0,08365 | 0,0725 | 0,6108 |
| 6° | 0,93163 | 0,08769 | 0,0665 | 0,6773 |
| 7° | 0,84395 | 0,01634 | 0,0603 | 0,7376 |
| 8° | 0,82761 | 0,16896 | 0,0591 | 0,7967 |
| 9° | 0,65864 | 0,06929 | 0,047 | 0,8438 |
| 10° | 0,58935 | 0,10097 | 0,0421 | 0,8859 |
| 11° | 0,48839 | 0,0327 | 0,0349 | 0,9207 |
| 12° | 0,45569 | 0,02301 | 0,0325 | 0,9533 |
| 13° | 0,43268 | 0,21136 | 0,0309 | 0,9842 |
| 14° | 0,22132 | | 0,0158 | 1 |

Fonte: Dados da pesquisa.

Figura 1: Gráfico do critério *scree*

Fonte: Dados da pesquisa.

Considerando os princípios da parcimônia e interpretabilidade, foram analisadas matrizes de cargas fatoriais para modelos com três, quatro e cinco fatores. Os modelos com três e cinco fatores não apresentaram boa interpretabilidade, por outro lado o modelo com quatro fatores, que explica 53,82% da variância dos dados, apresentou dimensões latentes mais claras e coerentes. O Quadro 3 apresenta a matriz de cargas fatoriais rotacionadas para quatro fatores, suas nomenclaturas e a comunalidade associada a cada variável. As áreas hachuradas representam cargas fatoriais superiores a 0,4, indicando a quais fatores a variável mais se correlaciona.

Quadro 3: Matriz de cargas fatoriais rotacionadas e denominação dos fatores

| Variável | Fator 1 | Fator 2 | Fator 3 | Fator 4 | Comunalidade |
|--------------|--|-----------------------|------------------------|----------------------------|--------------|
| | Tamanho da comunidade acadêmica e do espaço físico | Infraestrutura básica | Capital humano e etnia | Infraestrutura extraclasse | |
| n_alunos | 0,8004 | 0,1465 | 0,0929 | -0,0045 | 0,6707 |
| n_func | 0,9042 | 0,1196 | -0,0382 | 0,076 | 0,8392 |
| n_salas_util | 0,8606 | 0,0570 | -0,0043 | 0,0465 | 0,7461 |
| superior | 0,0780 | 0,0652 | 0,4170 | 0,3197 | 0,2864 |
| esp_mes_doc | 0,0773 | 0,1525 | 0,7237 | 0,0221 | 0,5535 |
| bbt | 0,1467 | 0,0759 | 0,0217 | 0,6358 | 0,4320 |
| lab | 0,0841 | -0,0597 | 0,1214 | 0,6224 | 0,4127 |
| net | 0,0682 | 0,2097 | -0,0783 | 0,4516 | 0,2587 |
| esgoto_rp | 0,1473 | 0,6802 | 0,2421 | -0,0554 | 0,5460 |
| lixo_cp | 0,0696 | 0,7574 | 0,1384 | 0,0402 | 0,5992 |
| agua_rp | 0,0915 | 0,8244 | 0,0221 | 0,0396 | 0,6900 |
| raca | -0,0981 | 0,1484 | 0,738 | 0,0627 | 0,5802 |
| resid | 0,4739 | 0,6368 | 0,0369 | 0,0605 | 0,6351 |
| nota_obj_red | 0,136 | 0,008 | 0,4302 | -0,2862 | 0,2856 |

Fonte: Dados da pesquisa.

Com relação às comunalidades, é recomendado que apresentem valor superior a 0,5, ou seja, que 50% da variância da variável seja explicada pelos fatores extraídos; mas as variáveis *superior*, *bbt*, *lab*, *net*, e *nota_obj_red* não cumprem esta recomendação. Quanto às variáveis *bbt*, *lab* e *net*, existe certa predisposição a apresentarem problemas de estimação por se tratarem de *dummies* (HAIR *et al.*, 2009). Quanto às variáveis *superior* e *nota_obj_red*, suas correlações com o fator 3 foram relativamente baixas, próximas a 0,4, o que indica baixa proporção de variância única explicada pelos fatores. As variáveis com baixa comunalidade foram hierarquicamente retiradas da análise, contudo, este procedimento não resultou em melhores indicadores de viabilidade e interpretabilidade dos fatores. Mantendo-se então essas variáveis, testou-se a opção com cinco fatores, o que elevaria as comunalidades, todavia esta alternativa tornou mais complexa a interpretação dos dados. Optou-se por manter a estimação com quatro fatores, tal como se apresenta no Quadro 2.

3.3 Caracterização e interpretação dos fatores

Tamanho da comunidade acadêmica e do espaço físico: o primeiro fator, que explicou 24,55% da variância total, está mais relacionado com as variáveis *n_alunos*, *n_func*, *_salas_util* e *resid*. As duas primeiras medem a quantidade de pessoal envolvido nas atividades da escola, enquanto *n_salas_util* mede o tamanho físico da mesma. A variável *resid* diz respeito ao percentual de alunos que residem em zona urbana e pode estar correlacionada ao fator por uma questão de densidade demográfica, ou seja, escolas da zona urbana que são maiores, têm mais alunos residentes em zona urbana, que por sua vez é mais densa que a zona rural.

Infraestrutura básica: o segundo fator, que explicou 12,70% da variância total, compreende as variáveis *esgoto_rp*, *lixo_cp*, *agua_rp* e *resid*. As três primeiras variáveis dizem respeito a condições mínimas de estrutura para o funcionamento da escola, a última se relaciona ao local de

residência dos alunos que a frequentam (urbano/rural). Escolas com melhor estrutura estão localizadas em zona urbana e por isso são aquelas com maior quantidade de alunos que residem em tal localização. Como mostrado anteriormente, dentre as escolas do meio rural, quase três quartos não possuem esgotamento sanitário por rede pública e quase a metade não usufrui de serviço de coleta periódica do lixo; dados como estes podem corroborar a correlação apresentada.

Capital humano e etnia: o fator 3, que explicou 8,61% da variância, está mais relacionado com as variáveis *superior*, *esp_mes_doc*, *raca* e *nota_obj_red*. As variáveis *superior*, *esp_mes_doc*, e *nota_obj_red* dizem respeito à capacitação do corpo docente da escola e ao desempenho dos alunos. O agrupamento destas variáveis em um único fator sugere relação estreita entre elas e pode ser indício de uma relação de causalidade, que poderia ser objeto de novos estudos. Uma vez que características docentes e discentes agruparam-se em um único fator, esperava-se que a carga fatorial da variável *nota_obj_red* em relação ao Fator 3 fosse mais relevante, diferentemente do valor observado de 0,4302, dado que a qualidade dos alunos deve ser diretamente ligada à qualidade dos professores. Esse pode ser um indício do argumento de Andrade e Soida (2013), que sugerem que o ENEM não é o meio mais adequado para medir o desempenho das escolas devido ao baixo número de alunos que realiza o exame em alguns casos.

O agrupamento da variável *raca* junto às *superior*, *esp_mes_doc* e *nota_obj_red* implica que alunos brancos e amarelos têm acesso a melhores professores e têm melhor desempenho no aprendizado. Este é um resultado que pode ser entendido de duas formas. Por um lado, isto pode ser reflexo da falta de variáveis de natureza socioeconômica

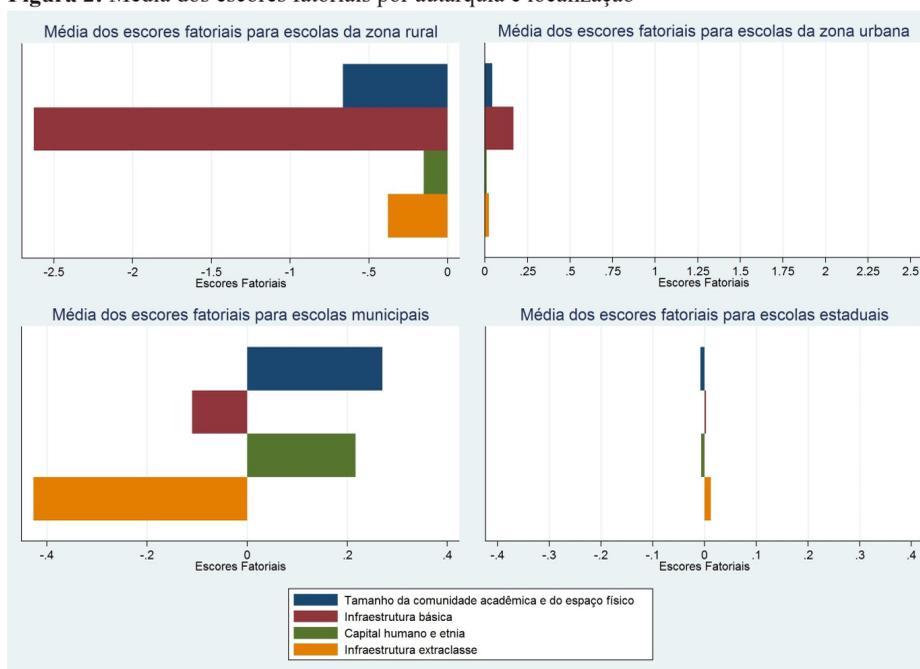
ou de escolaridade dos pais, que talvez pudessem ter maior correlação com a variável *raca* e formar um novo fator. Por outro lado, este agrupamento pode representar a realidade brasileira de segregação no ensino, em que negros e pardos têm, historicamente, menos acesso ao ensino de qualidade.

Infraestrutura extraclasse: o último fator, que explicou 7,96% da variância, se correlaciona de maneira mais forte com as variáveis *bbt*, *lab* e *net*. Essas variáveis indicam suporte à pesquisa e ao ensino, bem como a possibilidade de assimilar na prática o conhecimento teórico com atividades fora de sala de aula. Pode ainda ser entendido como o estado desejável de estrutura das escolas, que vai além do estágio de infraestrutura básica e condições mínimas para o funcionamento da escola (Fator 2).

Pode-se dizer que os fatores 2, 3 e 4 agrupam características sociogeográficas e de qualidade no ensino que podem implicar de maneira direta ou indireta na qualidade da formação dos alunos. O Fator 1 não apresenta essa propriedade, pois revela a dimensão latente de tamanho da escola e é plausível presumir que não necessariamente escolas maiores devam ser mais bem estruturadas e de melhor qualidade.

O Teste de Hotelling de diferença entre médias, aplicado aos escores fatoriais, rejeitou a 1% de significância a hipótese nula de que o vetor de médias fosse igual nas subamostras urbano/rural e municipal/estadual, o que corrobora a robustez da análise fatorial aplicada aos dados. A Figura 2 apresenta um esquema para a interpretação comparativa dos resultados de acordo com os escores fatoriais para autarquia e localização. Para a análise dos escores fatoriais, deve-se ressaltar a pressuposição metodológica de que, quando considerada toda a amostra, os escores possuem média zero.

Figura 2: Média dos escores fatoriais por autarquia e localização



Fonte: Dados da pesquisa.

Dessa forma, é natural que escolas urbanas e estaduais, cujas amostras são maiores (2047 e 2117, respectivamente), apresentem os escores mais próximos de zero. Os escores fatoriais estimados para escolas da zona rural são negativos e de maior magnitude em todos os casos, chegando a aproximadamente -2,6 para *infraestrutura básica*. Os escores de escolas urbanas são ligeiramente positivos, o que demonstra a melhor estruturação relativa das escolas da cidade frente às do campo. O fator menos discrepante entre subamostras de localização é *capital humano e etnia*, indicando que, apesar da carência de condições estruturais mínimas nas escolas do campo, a discrepância entre a qualidade dos professores e alunos para as escolas da cidade é relativamente pequena. Quanto às subamostras de autarquia, nota-se que escolas municipais são maiores, possuem melhor qualidade de capital humano, mas estão defasadas quanto à *infraestrutura extraclasse*. Por outro lado, as escolas estaduais apresentam escores mais próximos de zero e são mais bem qualificadas quanto às infraestruturas básica e extraclasse.

4 Conclusão

A pesquisa se propôs a avaliar, por meio da Análise Fatorial Exploratória, características das escolas municipais e estaduais de ensino médio de Minas Gerais, sendo consideradas análises por autarquia administrativa e localização rural/urbana. Os resultados da AFE confirmaram, em parte, a hipótese de que os dados se agrupariam em fatores referentes à estrutura da escola, corpo docente e corpo discente.

A estrutura fatorial encontrada agrupou características de corpo docente e discente em um único fator (Fator 3), dividiu as estruturais em dois fatores (Fatores 2 e 4) e instituiu um fator que se refere ao tamanho da escola e da comunidade acadêmica (Fator 1). Apesar de esses fatores explicarem pouco mais de 50% da variância dos dados, esta estrutura mostrou-se a mais clara quanto a sua interpretabilidade durante o processo de estimação.

A observação dos escores fatoriais apontou as deficiências relativas dentre cada uma das subamostras utilizadas. Notam-se maiores disparidades entre os meios rural e urbano que entre as autarquias municipal e estadual e, além disso, que estas disparidades são mais robustas na infraestrutura básica das escolas do que na formação dos professores e no desempenho dos alunos. Isto indica que atender a demandas de saneamento básico, tratamento de lixo e acesso à água nas escolas do campo deve ser prioridade na busca pela igualdade nas condições de ensino. Quanto à autarquia administrativa, as maiores discrepâncias residem na infraestrutura extraclasse, de modo que as escolas municipais estão, em média, aquém das estaduais. Essas escolas deveriam investir, portanto, na estruturação de laboratórios, acesso a internet e outras opções de estudos fora de sala. Constatou-se, ainda, que pode haver uma associação entre escolas com professores mais capacitados e maiores proporções de alunos brancos, o que pode representar um traço intrínseco de segregação no ensino,

em que alunos negros têm menores oportunidades. Neste sentido, poderiam ser aplicadas políticas de equalização do ensino por meio do oferecimento de oportunidades específicas a pardos e negros, tal como já é feito no caso das cotas em universidades públicas.

A inclusão de outras variáveis de características sociais, culturais e financeiras poderia aumentar o poder explicativo dos fatores e levar à construção de novas estruturas fatoriais, sendo esta uma sugestão para pesquisas futuras. Verificou-se também que a baixa correlação entre *nota_obj_red* e o Fator 3 pode indicar certa inadequação do uso do ENEM como *proxy* para a qualidade do ensino nas escolas. Apesar dessas limitações, acredita-se que o trabalho possa orientar outros esforços no sentido de avaliar a aplicabilidade da Análise Fatorial às características escolares, bem como auxiliar no entendimento da realidade e no desenvolvimento de ações que conduzam à igualdade e liberdade de acesso ao ciclo completo do ensino básico.

Referências

- ANDRADE, E.C.; SOIDA, I.A.I. *O ranking das escolas de Ensino Médio baseado no ENEM é confiável?*. Instituto de Ensino e Pesquisa. 2013. Disponível em: <http://www.insper.edu.br/wp-content/uploads/2012/12/2012_wpe283.pdf>. Acesso em: 2 abr. 2014.
- CARNEIRO, M.A. *LDB fácil: leitura crítico-compreensiva*, artigo a artigo. Petrópolis: Vozes, 2010.
- CATTANI, A.D. *Trabalho e tecnologia: dicionário crítico*. Belo Horizonte: Vozes. 1997.
- COSTA, G.L.M. *O ensino médio no Brasil: desafios à matrícula e ao trabalho docente*. Rev. Bras. Estudos Pedag., v.94 n.236 p.185-210, 2013.
- HAIR, J.F.; BLACK, W.C.; BABIN, B.J. *et al. Análise multivariada de dados*. Porto Alegre: Bookman, 2009.
- INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *A educação no Brasil rural*. Brasília: INEP, 2006.
- INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Panorama da educação do campo*. Brasília: INEP, 2007.
- INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sobre o ENEM*. Brasília: INEP, 2011a.
- INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados para download*. Brasília: INEP, 2011b.
- INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo Escolar: o que é o Censo Escolar?*. Brasília: INEP, 2011c.
- KUENZER, A.Z. O ensino médio no Plano Nacional de Educação 2011-2020: superando a década perdida? *Educ. Soc.*, v.31, n.112, p.851-873, 2010.
- LIMA, L.C.A. Da universalização do ensino fundamental ao desafio de democratizar o ensino médio em 2016: o que evidenciam as estatísticas? *Rev. Bras. Estudos Pedag.*, v.92, p.268-284, 2011.
- MÁRIO, P.C. *Análise discriminante*. In: CORRAR, L.J.; PAULO E.; DIAS FILHO, J.M. (Org.). *Análise multivariada para os*

cursos de administração, ciências contábeis e economia. São Paulo: Atlas, 2007. p.232-279.

MELO, S.D.G.; DUARTE, A.; Políticas para o ensino médio no brasil: perspectivas para a universalização. *Cad. Cedes*, v.31, n.84, p.231-251, 2011.

MINGOTI, S.A. *Análise de dados através de métodos de estatística multivariada: Uma abordagem aplicada*. Belo Horizonte: UFMG, 2005.

SIMÕES, C.A. *O ensino médio no Brasil: realidade e desafios*. In: SEMINÁRIO DE POLÍTICAS PÚBLICAS PARA O ENSINO MÉDIO. 2009. Belo Horizonte: Universidade Federal de Minas Gerais, 2009.

STATA CORP. *Stata 13 multivariate statistics*. College Station, TX. Stata Press, 2013. Disponível em: <<http://www.stata.com/manuals13/mv.pdf>>. Acesso em: 20 abr. 2015.