

Proposição e cálculo do Índice de Desenvolvimento Educacional (IDE) para os estados brasileiros: uma aplicação da análise fatorial de 2007 a 2011*

Dieison Lenon Casagrande**

Doutorando no Programa de Pós-Graduação (PPG) em Economia da Universidade Federal de Pernambuco (PIMES-UFPE) e Mestre pelo PPG em Economia e Desenvolvimento da Universidade Federal de Santa Maria (PPGE&D-UFSM)

Paulo Henrique de Oliveira Hoeckel***

Doutorando em Economia no Programa de Pós-Graduação em Economia da PUCRS e Mestre pelo PPGE&D-UFSM

Claílton Ataídes de Freitas****

Doutor em Economia Aplicada pela Universidade de São Paulo (USP) e Professor do Departamento de Ciências Econômicas da UFSM e do PPGE&D-UFSM

Cezar Augusto Pereira dos Santos*****

Professor da Unochapecó e Mestre pelo PPGE&D-UFSM

Resumo

O artigo tem como objetivo construir um Índice de Desenvolvimento Educacional (IDE) para elencar os 27 estados brasileiros, durante o período de 2007 a 2011, através do ferramental dos Métodos Quantitativos da Estatística Multivariada, pela aplicação da metodologia da análise fatorial (AF), atra-

* Artigo recebido em ago. 2013 e aceito para publicação em out. 2015.

 Este trabalho está licenciado sob uma Licença Creative Commons Atribuição-NãoComercial 4.0 Internacional.

Open Acces (Acesso Aberto)

Revisão de Língua Portuguesa: Breno Camargo Serafini

** E-mail: dieisonlenon@yahoo.com.br

*** E-mail: ph.hoeckel@gmail.com

**** E-mail: caf@ccsh.ufsm.br

***** E-mail: cezarsantos1975@hotmail.com

vés da qual serão adotados os escores fatoriais para a elaboração do referido índice. Seu uso baseia-se no interesse em contar com um índice único para cada estado do Brasil, para cada período, e, assim, hierarquizar as unidades da Federação. Os resultados encontrados para o IDE evidenciam certa heterogeneidade regional. Assim, o IDE médio para o período é de 0,371, e o Distrito Federal e os Estados de São Paulo e Rio de Janeiro encontram-se no topo desse *ranking*. Por outro lado, os estados da Região Nordeste, como Piauí, Maranhão, Bahia e Paraíba, encontram-se na parte inferior dessa hierarquização.

Palavras-chave

Índice de Desenvolvimento Educacional; análise fatorial; estados brasileiros

Abstract

This paper aims to build an Educational Development Index (EDI) to rank the 27 Brazilian states during the period from 2007 to 2011, through the tools of Quantitative Methods of Multivariate Statistics, by applying the methodology of Factorial Analysis, which will provide the factorial scores for the preparation of this index. Its use is based on the interest in having a single index for each Brazilian state for each period, thus establishing a hierarchy of the country. The results obtained for the EDI show some regional heterogeneity. So the average of the EDI for the period is 0.371, and the states of Distrito Federal, São Paulo and Rio de Janeiro are at the top of this ranking. By contrast, the states of the Northeast, such as Piauí, Maranhão, Bahia and Paraíba, are at the bottom of the ranking.

Keywords

Educational Development Index; factorial analysis; Brazilian States

Classificação JEL: C02, I21

1 Introdução

Transformações sociais, mudanças e progressos tecnológicos e a globalização estão entre as principais causas da transformação e da evolução do sistema educacional brasileiro, sendo esse um campo amplamente discutido na literatura¹. A educação necessita de mudanças, para que possa se adequar às demandas da sociedade contemporânea, visando atingir suas expectativas e objetivos quanto ao mercado de trabalho e quanto à minimização dos problemas sociais. Assim, espera-se que o sistema educacional brasileiro, quando avaliado, principalmente por órgãos de caráter internacional, perca o rótulo de ser apontado como um dos fatores responsáveis pelo baixo grau de desenvolvimento do País.

Segundo Menezes Filho (2001), a importância da educação para o bem-estar de uma nação já foi bastante documentada. Muitos estudos apontam o investimento em capital humano como sendo o responsável por grande parte das diferenças de produtividade entre os países.² Dessa forma, em nível mundial, muitos estudos têm sido realizados, com o intuito de investigar os efeitos da educação sobre o crescimento econômico, e encontram fortes evidências dessa relação. Dentre os mais recentes, cabe destacar: Lee e Lee (1995), Barro (2001), Bosworth e Collins (2003), Stevens e Weale (2003), Podrecca e Carmeci (2004), Jamison, Jamison e Hanushek (2007), Altinok (2007), Hanushek e Woessmann (2008, 2009), Aghion *et al.* (2009) e Sahlgren (2014).

Historicamente, a questão da educação não fora tratada como a prioridade dos governos. Contextualizando esse argumento, da década de 50 a meados dos anos 80, o principal objetivo dos planos de governo era a promoção da industrialização (claro que existiam objetivos secundários) e, como consequência, o aumento do endividamento externo (mesmo durante o período do Milagre Econômico e durante a década de 70). Ao longo das décadas de 80 e 90, a prioridade central foi o combate à inflação e ao baixo crescimento econômico (da mesma forma, associados a outros fatores, como a crise da dívida na primeira metade da década de 80), e, nos anos 2000, a prioridade se voltou para o crescimento econômico e, mais recentemente, para o combate aos efeitos da crise mundial. Portanto, com esse argumento, percebe-se que os problemas de cunho estritamente econômico predominaram (ou predominam) nas políticas governamentais, e a educação ainda não foi percebida pelos governos como fundamental para garantir

¹ Autores como Oliveira (2009), Sander (2007), Gadotti (2000) e Dourado (2007), dentre outros, discutem a educação e seu processo de evolução na atualidade.

² Para mais informações, ver Hall e Jones (1998) e Mankiw, Romer e Weil (1992).

o crescimento sustentável e alavancar o desenvolvimento econômico. Conforme Castro (2009), em países desenvolvidos, a educação é parte das políticas sociais e compõe o núcleo do sistema de promoção social, mediante a sua capacidade de ampliar as oportunidades para os indivíduos, além de ser um elemento estratégico para o desenvolvimento econômico.³

Apesar de progressos limitados entre 1999 e 2007 (a partir de 2008, houve certa estagnação na melhora desses resultados, principalmente devido à crise econômica), o Brasil continua sendo a nação com a maior população de crianças fora da escola na região do Caribe e América Latina e o décimo segundo país na esfera mundial, com cerca de 3,36 milhões de pessoas de quatro a 17 anos fora da escola, em 2013.⁴ Além disso, o que é repassado nas escolas coloca o estudante brasileiro como um dos mais despreparados no conhecimento de língua pátria, matemática e ciência, na amostra realizada em 160 países pela UNESCO. Isso significa que, em média, os estudantes brasileiros têm dificuldade de se expressar, quer seja na linguagem falada, quer seja na escrita, e conhecem pouco os fundamentos básicos da matemática. A qualidade da educação oferecida às crianças brasileiras é outro aspecto considerado nos relatórios da UNESCO. Conforme o relatório de 2010 (UNITED NATIONS EDUCATIONAL SCIENTIFIC AND CULTURAL ORGANIZATION, 2010), no Brasil, assim como no Chile, na Colômbia, no México, no Uruguai e na Argentina, de 36% a 58% dos alunos foram incapazes de demonstrar a habilidade de leitura normalmente alcançada no meio do período da escola primária, em países desenvolvidos.

Assim, é de primordial importância e relevância para o crescimento e desenvolvimento econômico do Brasil que, em todos os estados, o acesso à educação seja democrático e que o ensino público seja tão bom, se não melhor, do que o ensino das escolas privadas. Isso é uma realidade em termos de ensino superior, mas é imprescindível que o seja, também, para os ensinos fundamental e médio (HADDAD, 2003).

Educação e crescimento estão diretamente interligados no que diz respeito à formação de profissionais qualificados e, conseqüentemente, no aumento do nível de produção do País. Além do mais, a educação representa um valor intrínseco para o desenvolvimento de um país, pois altos níveis de escolaridade aumentam os rendimentos das pessoas, promovem

³ De acordo com Fernandes e Gremaud (2009), a questão do acesso à escola mereceu, nos anos 80 e 90, um certo esforço, sendo que o País conseguiu colocar quase a totalidade das crianças na escola. Porém, em relação aos indicadores de fluxo escolar, o Brasil ainda permanece com fortes dificuldades: altas taxas de repetência e elevada proporção de adolescentes que abandonam a escola sem concluir a educação básica; isso aliado à baixa proficiência obtida em exames padronizados.

⁴ Informações referentes ao Relatório de Monitoramento Global de EPT, 2013 (UNITED NATIONS EDUCATIONAL SCIENTIFIC AND CULTURAL ORGANIZATION, 2013).

desenvolvimento social, inserem o país em um patamar mais produtivo e contribuem para a redução de desigualdades inter-regionais.

Segundo Barros *et al.* (2001), no Brasil, a combinação de um sistema educacional público precário com as graves imperfeições do mercado de crédito tem feito com que o nível de investimentos em capital humano esteja sistematicamente abaixo dos padrões internacionais. Esse fato surpreende, na medida em que todas as estimativas existentes para as taxas de retorno desse tipo de investimento apresentam valores bastante atraentes.

Ainda de acordo com Barros *et al.* (2001), a despeito das elevadas taxas de retorno à educação, o Brasil apresenta indicadores educacionais abaixo dos padrões internacionais. Ainda mais preocupante é o fato de que o subinvestimento em capital humano é tanto mais acentuado quanto mais pobre é a família. Dado que pessoas menos escolarizadas serão, com maior probabilidade, pobres no futuro, essa natureza diferenciada do investimento em educação leva à transmissão intergeracional da pobreza, ou seja, os indivíduos nascidos em famílias pobres hoje tenderão a ter escolaridade inferior e serão, com maior probabilidade, os pobres de amanhã. E assim, o reflexo desses aspectos é sentido no mercado de trabalho, principalmente com os baixos índices de produtividade apresentados na economia brasileira.

Para Souza e Oliveira (2006), a ampliação da expansão do estoque de conhecimentos gera um aumento mais que proporcional, nos níveis de renda. Ainda segundo os autores, o capital humano, o capital físico e a força de trabalho estão intimamente associados ao conhecimento técnico. Desse modo, os novos conhecimentos produzem externalidades positivas, que são apropriadas pelos agentes produtivos, elevando o nível de produção agregada, e, assim, as regiões que mantiveram investimentos crescentes na ciência básica e aplicada, na descoberta de novos produtos e processos e em educação e saúde de sua população tendem a crescer mais rapidamente.

Como é estreita a relação da educação com as outras várias dimensões sociais (saúde, redução do crescimento populacional, queda da mortalidade infantil, aumento da expectativa de vida) e econômicas (renda *per capita*, crescimento econômico, distribuição de renda), então, ao se propor construir um índice que retrate a situação da educação em nível das unidades da Federação no Brasil, não se pode negligenciar o seu caráter multidimensional⁵.

Diante dessa problemática, o presente estudo tem por objetivo construir um índice, ou indicador, que contemple diferentes aspectos do desenvolvimento da educação no Brasil, através do ferramental estatístico da análise fatorial (AF) e, a partir deste, diagnosticar, em termos relativos, a situação

⁵ O conceito de multidimensional está associado ao amplo caráter de agrupamento físico em um único indicador de várias dimensões, simultaneamente.

da educação em cada estado brasileiro. É de comum acordo na literatura que um indicador de desenvolvimento educacional pode ser empregado com sucesso em comparações entre diferentes espaços geográficos, o que facilita o desenvolvimento de políticas específicas para enfrentar as necessidades individuais.

Para atender a tais objetivos, o presente artigo encontra-se dividido em seis seções. A primeira contempla esta breve **Introdução**, enquanto a segunda seção apresenta a justificativa e motivação do estudo. Na terceira, traz-se uma abordagem de alguns dos principais indicadores educacionais presentes na literatura. Na sequência, as quarta e quinta seções abordam, respectivamente, a metodologia e os resultados da pesquisa. Por fim, finaliza-se com a seção de **Conclusões**.

2 Motivação e justificativa

O diagnóstico da situação relativa da educação em nível regional contribui com a discussão sobre qual a melhor estratégia de desenvolvimento (e/ou crescimento) para o País, para, assim, livrá-lo desse estágio não confortável em que se encontra a educação, facilitando, dessa forma, a elaboração de políticas específicas para enfrentar os problemas estruturais da educação no Brasil.

A justificativa para a elaboração desse índice ocorre à luz de dois pontos centrais. O primeiro diz respeito à sua importância como mecanismo de elaboração de políticas públicas. Além da existência dos já tradicionais índices de mensuração das diferenças regionais — como o Índice de Desenvolvimento Humano (IDH), por exemplo —, a construção desse indicador propõe um mecanismo prático e com um foco central: a mensuração da dimensão educação. Esse aspecto representa uma das principais causas das diferenças interestaduais e pode vir a servir como um expoente no direcionamento das principais medidas a serem adotadas pelos formuladores de políticas na área de educação, com o intuito da redução desse *gap* educacional entre as regiões, para, assim, o País crescer de forma mais homogênea.

Quanto ao segundo ponto, a essência do método de análise proposto (análise fatorial), por ser um método *data-driven* e não apresentar o critério de escolha de variáveis delimitado pela literatura precedente, possibilita a extração de informações, expressas realmente pelo comportamento das variáveis em análise. A abordagem *data-driven* não necessita de modelos *ad hoc*, evitando-se, assim, o direcionamento dos resultados por critérios subjetivos. Em contrapartida, tal direcionamento será dado pela qualidade e

pela quantidade de dados disponíveis, daí a necessidade de uma ampla gama de dados para a confiabilidade dos resultados provindos dessa análise.

Logo, o principal motivo pela escolha do método proposto está no fato de querer avaliar o efeito real das variáveis na construção do índice, evitando a indução predeterminada da importância de cada dimensão na sua essência, como, por exemplo, o IDH, que apresenta escores iguais em sua construção para dimensões que não necessariamente têm o mesmo impacto, feita de forma *ad hoc*.

Dado que a dimensão educação é influenciada por diversas variáveis⁶, e que diferentes indicadores já foram construídos para a sua avaliação, mensurá-la se torna um desafio, dado que esta não pode ser diretamente observada. Segundo Blalock (1974), ainda que o desenvolvimento da teoria seja importante em si mesmo, os mais sérios e importantes problemas que requerem imediata e forte atenção são aqueles de conceitos e mensuração.

Nesse sentido, Zeller e Carmines (1980) afirmam que a transformação de conceitos em indicadores empíricos não é um processo simples, pois o mesmo conceito pode ser operacionalizado de formas diferentes. Dessa forma, conforme Tabachnick e Fidell (2007), um processo alternativo de mensuração é identificar variáveis que “caminham juntas”, ou seja, variáveis que apresentam a mesma estrutura subjacente. Tecnicamente, isso pode ser implementado através da análise fatorial, cuja principal função, das diferentes técnicas de análise fatorial, é reduzir uma grande quantidade de variáveis observadas a um número reduzido de fatores.

No entanto, nessa perspectiva, para Freitas, Paz e Nicola (2007), a análise fatorial é suscetível de críticas, principalmente pelo fato de que, ao selecionar as relações mais importantes, ajuda a interpretar as relações que surgem de cada fator separado. Como as escolhas e as interpretações são, em maior ou menor medida, subjetivas, não se pode assegurar que essas relações sejam as únicas e verdadeiras. Apesar dessa crítica, o método da análise fatorial é uma ferramenta importante para a definição de um padrão de relações específico⁷.

Agora, para concretizar essa junção do método de abordagem com nosso problema de pesquisa, cabe uma maior argumentação acerca dos fundamentos perante os elementos de avaliação da educação. Segundo

⁶ Ver Glewwe *et al.* (2011), que analisaram mais de 9.000 estudos publicados entre 1990 e 2010, tanto na literatura da educação quanto na literatura econômica, para investigar quais características da escola e específicas dos professores têm fortes impactos positivos na aprendizagem e no tempo de permanência do aluno na escola.

⁷ Dentre outras desvantagens da utilização desse método, podemos citar: a nomenclatura dos fatores pode ser difícil, sendo que vários atributos podem estar altamente correlacionados, mas sem nenhuma razão aparente; sendo completamente independentes as variáveis observadas, a análise fatorial é incapaz de produzir um padrão significativo.

Fernandes e Gremaud (2009), existem hoje evidências robustas de que a escolaridade, medida como anos de estudo completos, é uma variável-chave na determinação do progresso econômico de indivíduos e nações. Além dos aspectos econômicos, a escolaridade tem sido associada também a uma variedade de benefícios não econômicos: melhor saúde, redução da criminalidade, menor incidência de gravidez na adolescência, maior coesão social, dentre outros. Enquanto os benefícios da educação são bem conhecidos, a maneira pela qual ela gera esses benefícios é ainda pouco compreendida. Nesse aspecto, a hipótese mais difundida é que, na escola, os indivíduos adquirem determinados conhecimentos e desenvolvem habilidades que possuem um alto valor social.

De acordo com Fernandes e Gremaud (2009), as avaliações educacionais podem ser vistas como um caso particular das avaliações de programas e/ou políticas e, desse modo, estão relacionadas à ideia de resultados. Independentemente do contexto em que elas estão inseridas e de seus objetivos, as avaliações educacionais, ou buscam aferir resultados passados, ou inferir resultados futuros. Isso não significa que elas sejam sempre conduzidas com base em indicadores de resultados. Além disso, o debate sobre avaliações orientadas por resultados e avaliações orientadas por insumos e processos é, em grande parte, um debate sobre medida. Caso duas medidas, uma baseada em resultados e outra baseada em insumos e processos, convergissem, no sentido de apresentarem a mesma ordenação de escolas, a medida baseada em insumos e processos deveria ser preferida. Isto porque ela dá mais retorno às escolas, pois, além de dar um diagnóstico da qualidade, sinaliza quais os pontos mais vulneráveis que deveriam ser atacados (FERNANDES; GREMAUD, 2009).

No entanto, ainda conforme Fernandes e Gremaud (2009), diversos estudos têm apontado a fraca correlação (condicional nas características observáveis dos estudantes) entre as variáveis de insumos e processos e o desempenho dos estudantes em exames padronizados. Essa fraca correlação tem sido a base para as críticas de medidas de qualidade da escola baseadas em insumos e processos, pois, como ressaltam Hanushek e Raymond (2003), uma recompensa para as escolas pode advir da melhora de processos, ao invés de melhora no desempenho dos estudantes. Esse tipo de argumentação, no entanto, pressupõe que essa fraca correlação decorre da deficiência das variáveis de insumos e processos em explicar a qualidade da escola, e não da debilidade dos resultados dos testes em medir essa qualidade.

Contudo Fernandes e Gremaud (2009) afirmam que, embora boa parte dos problemas levantados já estivesse presente há muito tempo, nas avaliações educacionais, eles não costumavam criar tanta polêmica como nos

dias de hoje. A polêmica atual está ligada à nova função que os exames assumem nas políticas de *accountability*⁸, cujo cerne é considerar não apenas os alunos, mas escolas, professores, diretores e gestores como responsáveis pelo desempenho dos estudantes.

Porém, como ressaltado por Fernandes e Gremaud (2009), tais programas buscam verificar se a aprendizagem proporcionou, aos estudantes, determinadas competências e conhecimentos. É importante ressaltar que os programas de *accountability* surgem nos países desenvolvidos, onde o problema de acesso à escola e permanência nela está praticamente resolvido. A defasagem idade-série não é também um problema sério nesses países. Em países onde esses problemas não estão resolvidos, enfatizar os resultados dos exames pode ser precipitado. Para Chay, McEwan e Urquiola (2005), o fenômeno de “reversão para a média” faz com que o desempenho médio das escolas em testes padronizados, como o Sistema de Avaliação da Educação Básica (SAEB), a Prova Brasil ou o Programme for International Student Assessment (PISA) possa variar muito de um momento de avaliação para outro.⁹

3 Aspectos educacionais do Brasil

No nível nacional, as diferenças educacionais na população são grandes, bem como os diferenciais salariais associados a essas diferenças. Tais diferenças regionais, fato bem comum à realidade brasileira, afloram, quando são analisados alguns números referentes ao desempenho estadual da educação. Apesar de o Brasil ter caminhado rumo à universalização do ensino, segundo Oliveira (2007), o único indicador que fora praticamente universalizado até o início deste século fora o acesso ao ensino fundamental, enquanto indicadores como repetência, abandono escolar, distorção idade/série e, principalmente, indicadores de aprendizagem, embora apresentem certo sucesso, desenharam um quadro de profundas desigualdades regionais. Ainda segundo Oliveira (2007), mesmo com a regularização do fluxo escolar, é comum encontrar falta de escolas em regiões específicas, o que se atribui ao fato de a oferta excedente não se encontrar em áreas onde as crianças são mais excluídas.

⁸ O termo *accountability* tem sido traduzido como transparência, responsabilização, prestação de contas e outros. Na falta de concordância sobre a melhor tradução, manteremos o termo em inglês.

⁹ Paro (2000) também faz uma crítica aos objetivos parciais das avaliações baseadas em exames padronizados.

Nessa mesma ótica, temos que um aspecto particularmente importante do sistema de ensino brasileiro é que, virtualmente, todos entram na escola; no entanto, apenas 84% concluem o quarto ano, 57% concluem o ensino fundamental, e apenas 37% atingem o nível de conclusão do ensino médio, dos quais apenas 28% saem com o devido diploma (INSTITUTO DE PESQUISA ECONOMICA E SOCIAL, 2006).

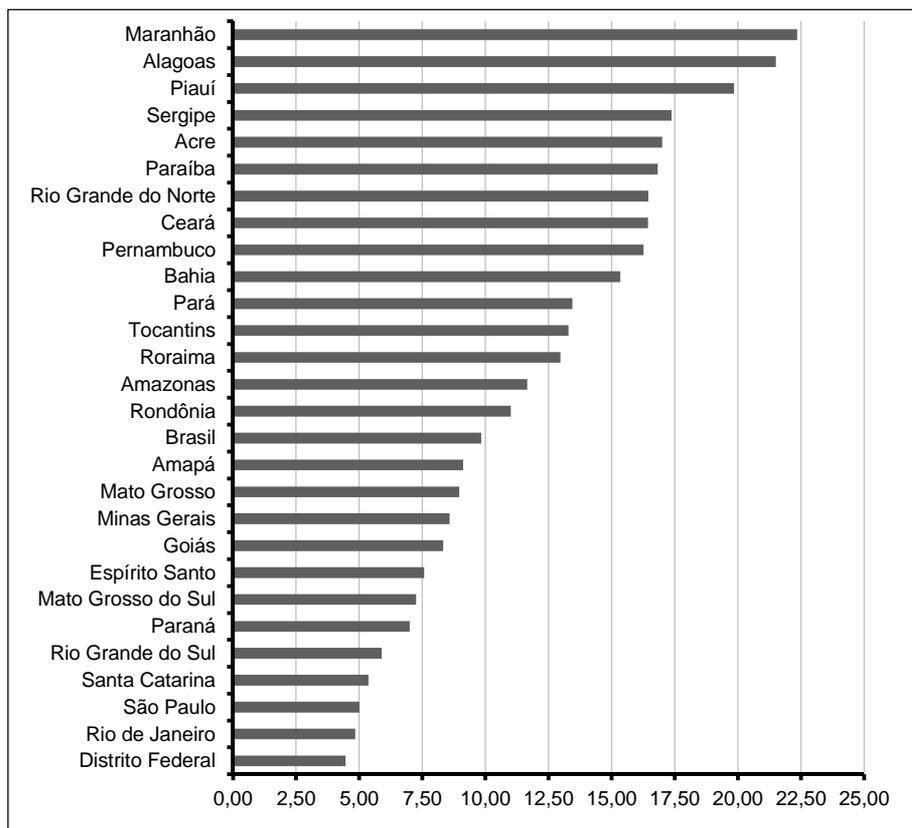
Uma análise histórica dos principais indicadores de desempenho educacional regional aponta que a sua evolução não se deu de forma convergente; pelo contrário, indica um aumento cada vez maior das desigualdades regionais.

Frequentemente, a taxa de analfabetismo é apontada como uma das principais evidências do atraso educacional do Brasil. A Figura 1 retrata o analfabetismo em cada unidade da Federação (UF) e no Brasil. Percebe-se uma grande heterogeneidade regional, ou seja, os estados das Regiões Norte e Nordeste, em sua quase totalidade, ficam acima da média nacional, enquanto os das demais regiões apresentam desempenho superior ao da média nacional. Esse desempenho heterogêneo fica mais evidente, ao se analisarem os extremos: enquanto o Distrito Federal aponta uma taxa de analfabetismo de 4,46%, Maranhão e Alagoas apresentam-na superior a 20%.

Apesar dos problemas estruturais da educação no Brasil, em pelo menos um quesito — número médio de anos de estudo — ela vem apresentando números favoráveis a partir de 2005. Isso está caracterizado na Figura 2, a qual revela uma evolução de quase 10% na população com maior nível de instrução, pois, em 2011, a população brasileira com mais de 11 anos de estudo representava 35,05% do total, contra os 27,12% apontados em 2005.

Figura 1

Percentual da taxa de analfabetismo das pessoas de cinco anos ou mais de idade, por unidade da Federação e no Brasil — 2011

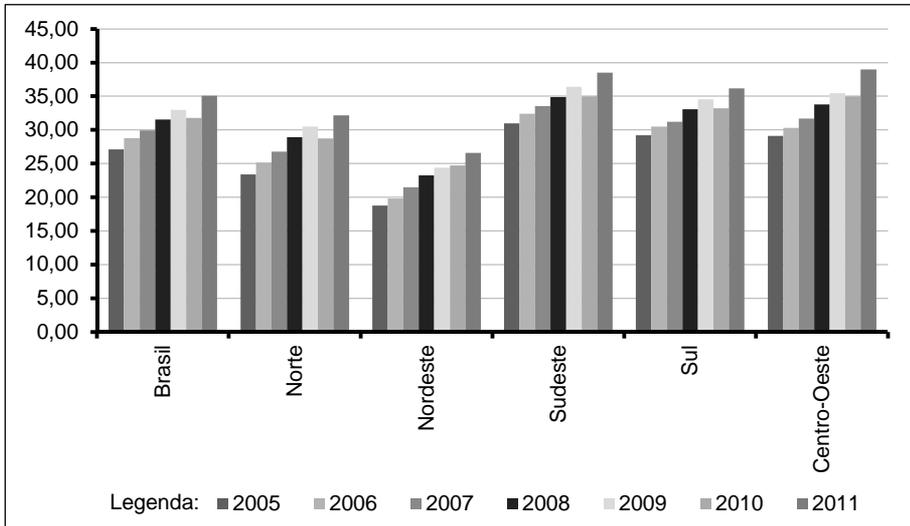


FORNE DOS DADOS BRUTOS: Censo da Educação do INEP (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2011).

Contudo persiste ainda uma grande diferença regional. Em 2011, a Região Centro-Oeste apresentou um percentual de 39% da população com mais de 11 anos de estudo, enquanto, na Nordeste, esse número reduziu-se para 26,5%. No entanto, cabe ressaltar que, nesse período, foi a Região Nordeste a com maior taxa de crescimento da população com mais de 11 anos de estudo, com cerca de 6,0%, enquanto a média nacional foi de 4,4%. Em termos gerais, houve expansão da qualificação pessoal, especialmente nas Regiões Norte e Nordeste, fato esse fundamental para uma maior geração de renda regional e, por consequência, fundamental para o crescimento regional.

Figura 2

Percentual de pessoas de 10 anos ou mais, com mais de 11 anos de estudo, por regiões e no Brasil — 2005-11

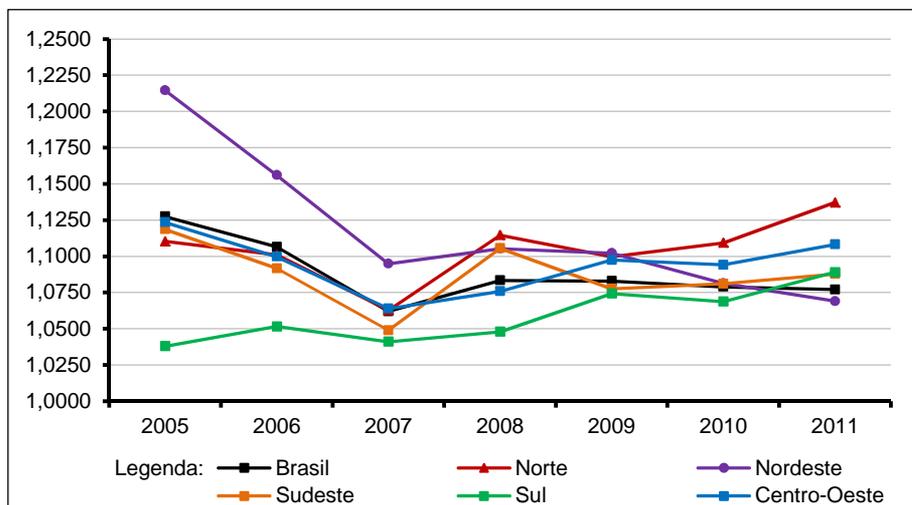


FORNTE DOS DADOS BRUTOS: Censo da Educação do INEP (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2005, 2006, 2007, 2008, 2009, 2010, 2011).

Outro indicador considerado que expressa o desempenho educacional brasileiro é a taxa de escolarização bruta — o qual demonstra o número de matrículas em determinado nível de ensino em relação à demanda potencial adequada para aquele nível. De acordo com a Figura 3, pode-se inferir que uma taxa de escolarização maior que 100% indica uma defasagem em relação a idade e conclusão do nível de ensino, fato atribuído tanto pela repetência, quanto pelo retorno de pessoas que não concluíram essa etapa. A análise desse indicador é um pouco dúbia, pois a sua variação pode ser reflexo de aspectos tanto negativos quanto positivos. Um fator positivo é dado pelo aumento do atendimento escolar, que eleva a taxa de escolarização bruta; por outro lado, essa taxa também pode aumentar, em decorrência do número de repetências nesse nível escolar. Em 2011, o Brasil atingiu uma taxa de 107,7%, o que indica que existia 7,7% a mais de pessoas cursando o ensino fundamental em relação ao universo da população de sete a 14 anos, ou seja, 7,7% das pessoas dessa faixa etária estão cursando um nível de ensino inadequado para a sua idade.

Figura 3

Taxa de escolarização bruta no ensino fundamental,
nas regiões brasileiras e no Brasil — 2005-2011



FONTE DOS DADOS BRUTOS: Censo da Educação do INEP (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2005, 2006, 2007, 2008, 2009, 2010, 2011).

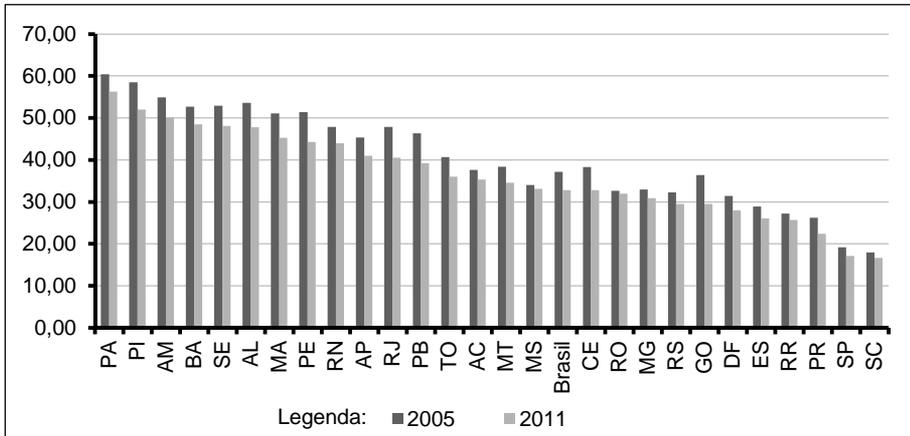
NOTA: No ano de 2011, a faixa etária referente ao ensino fundamental é de seis a 14 anos.

Por fim, analisar-se-á outro indicador que caracteriza o desempenho educacional no Brasil: a evolução do padrão de distorção idade/série, ou interrupção no fluxo escolar, o qual é considerado um dos maiores gargalos no sistema educacional brasileiro.

Verifica-se, na Figura 4, que essa taxa, em todos os estados brasileiros, declinou de 2005 a 2011. Contudo, para a maioria dos estados da Região Nordeste, essa taxa permaneceu acima de 40% em 2011, ou seja, em torno de 40% da população com idade ideal para cursar esse nível de ensino não o faz. O pior desempenho é verificado no Estado do Pará, onde essa taxa chegou a mais de 56%, em 2011. No outro extremo, está Santa Catarina, com 16,30%. Apesar da melhora ocorrida nesse período, ainda há muito a ser feito, dado que o ideal seria não haver interrupção no fluxo escolar.

Figura 4

Percentual da taxa de distorção idade/série no ensino médio,
nos estados e no Brasil — 2005 e 2011



FONTE DOS DADOS BRUTOS: Censo da Educação do INEP (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2005, 2006, 2007, 2008, 2009, 2010, 2011).

Nesta seção, foram abordados alguns dos indicadores que expressam o desempenho do sistema educacional brasileiro e evidenciam certas heterogeneidades regionais. No entanto, existe uma gama de outros indicadores que poderiam ser utilizados e que expressam, com a mesma importância, as diferenças regionais apontadas, como indicadores de rendimento escolar (taxa de aprovação/reprovação), razão aluno professor (tanto na educação básica quanto no ensino superior), qualificação dos professores (dado pelo nível de formação), gastos públicos com educação, dentre outros. Contudo foram selecionadas pelos autores as informações estatísticas que apontassem, de forma mais clara, essas diferenças.

4 Aspectos metodológicos

Nesta seção, inicialmente, são abordadas as variáveis utilizadas na análise. Num segundo momento, abordam-se os aspectos metodológicos da análise fatorial, e, por fim, apresenta-se o método de construção, através dos escores fatoriais, do Índice de Desenvolvimento Educacional.

4.1 Seleção das variáveis e fonte dos dados

A fim de mensurar o desempenho educacional no nível estadual, no Brasil, foram selecionadas variáveis que permitissem abranger vários aspectos, a partir das estatísticas disponíveis, ao longo do período 2007-11.

A pré-seleção das variáveis dá-se pelo referencial teórico da Economia da Educação, que aponta os principais determinantes de cada uma das dimensões elencadas. Nesse contexto, a recente literatura de Hanushek e Woessmann (2007, 2008, 2010, 2012) aponta o papel das habilidades cognitivas na promoção do desenvolvimento econômico, através do impacto da educação sobre os demais determinantes econômicos¹⁰. Outro aspecto determinante na definição das variáveis foi a sua disponibilidade ao longo do período de amostra.

Hanushek e Woessmann (2010) argumentam que o resultado econômico é determinado pelo nível de capital humano e por uma variedade de outros fatores. Sendo assim, a literatura aponta uma diversidade de variáveis (fatores) que afetam o desempenho da educação, bem como diversas são as medidas adotadas para avaliar o desempenho da educação. Entre os estudos que identificam os fatores que afetam a qualidade do ensino tem-se: Afonso e Aubyn (2006) analisam o número de professores por estudante e o tempo gasto na escola; Gundlach, Woessmann e Gmelin (2001) analisam os gastos por estudante; Hanushek e Luque (2003) investigam os efeitos do tamanho da classe e características relacionadas aos professores (utilizam como variáveis o tamanho da classe e a experiência e o grau de educação dos professores); Achilles (1999), Woessmann e West (2006) e Schanzenbach (2014) investigam o efeito do tamanho da classe sobre o aprendizado dos alunos; Campbell, Hombro e Mazzeo (2000), Barro e Lee (2001) e Organisation for Economic Co-Operation and Development (2014) investigam o efeito da relação aluno/professor e encontram um efeito positivo da menor relação aluno/professor sobre o aprendizado.

Ainda Hanushek e Woessmann (2007), Barber e Mourshed (2007) e Spaul (2013) indicam que a qualidade dos professores tem um impacto poderoso sobre o resultado dos alunos (justificando a inclusão de variáveis como número de professores doutores e mestres por instituição e número de habitantes e número de professores doutores na educação básica). Estudos como os de Barro (1991), Mankiw, Romer e Weil (1992) e Levine e Renelt (1992) utilizam a taxa de escolarização como uma medida de desempenho da educação. Além disso, estudos mais recentes, como Sala-i-

¹⁰ Essa literatura se desenvolve através da análise de microdados (no nível de indivíduos), análise esta que está além do escopo desta pesquisa.

-Martín, Doppelhofer e Miller (2004) e Baldacci *et al.* (2008) apontam que taxas de escolarização (de diversos níveis) são um dos determinantes mais robustos do crescimento econômico, com certa sensibilidade dos indicadores, associados aos modelos estatísticos utilizados. Por fim, Cascio, Clark e Gordon (2008) utilizam uma tradicional medida de escolaridade, anos de estudo.

Portanto, fundamentado na literatura citada e com base na disponibilidade informacional, foram selecionados 17 indicadores, que se julga serem capazes de captar e expressar essas diferenças estaduais. Esses indicadores contemplam diferentes categorias, como desempenho, capital humano, estrutura e fatores econômicos. Tais indicadores são apresentadas no Quadro 1.

Os dados utilizados para a execução da pesquisa são secundários e têm como fonte principal as sinopses estatísticas da educação básica e da educação superior do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP)¹¹. As informações referentes à população por classe etária provêm do Instituto de Pesquisa Econômica Aplicada (2012), e as referentes aos gastos com educação estão disponíveis no Sistema de Informações sobre Orçamentos Públicos em Educação (FUNDO NACIONAL DE DESENVOLVIMENTO DA EDUCAÇÃO, 2012).

Quadro 1

Descrição das variáveis

INDICADORES	DESCRIÇÃO	
TANALF	Taxa de analfabetismo (%). Representa o quociente entre a população analfabeta e a população total de um mesmo grupo etário. Considera-se analfabeto aquele indivíduo que é incapaz de ler e escrever ao menos um bilhete simples na sua língua de origem.	
DISTEF	Distorção idade/série - ensino fundamental	Corresponde ao total de matrículas de pessoas que estão cursando determinada série em idade superior à considerada ideal sobre o total de matrículas na série em questão. Esse índice é importante, por determinar problemas relacionados com a alta repetência em determinada série, que é um dos principais problemas no sistema de ensino brasileiro, com graves consequências para os níveis de escolaridade da população.
DISTEM	Distorção idade/série - ensino médio	

(continua)

¹¹ As sinopses estatísticas da educação básica e da educação superior do INEP são referentes ao período 2007-11. As variáveis selecionadas dessa base de dados são as seguintes: taxa de analfabetismo, taxa de distorção idade/série do ensino fundamental e do ensino médio, taxa de escolarização bruta do ensino fundamental, número de professores da educação básica, alunos matriculados, número de turmas, número de instituições de ensino, taxa de aprovação do ensino médio, matrículas e professores no ensino superior, qualificação dos profissionais (mestres e doutores), número de concluintes do ensino superior e o número de docentes com ensino superior.

Quadro 1

Descrição das variáveis

INDICADORES	DESCRIÇÃO
TESCBEF	Taxa de escolarização bruta no ensino fundamental (%). É dada pela razão entre as matrículas em um determinado nível de ensino e a população em idade adequada de cursar tal nível. Possibilita avaliar o volume de matrículas nesse nível, em função da demanda potencial na faixa etária adequada.
TRENDESC	Taxa de rendimento escolar (Taxa de aprovação no ensino médio - %). Corresponde ao total de alunos aprovados em determinado nível de ensino sobre o total de matrículas efetuado nesse nível de ensino.
COMA	Relação concluintes/matriculados no ensino superior (%). Corresponde ao total de alunos concluindo o ensino superior num determinado período sobre o total de matrículas.
PROFEB	Razão professor da educação básica/mil habitantes. Relaciona o número de professores com a população em geral. É um importante indicador, que demonstra a disponibilidade da oferta de professores para as unidades geográficas.
ANES	Anos de estudo (pessoas de 10 anos ou mais com mais de 11 anos de estudo - %). Corresponde à média dos anos de estudos concluídos por uma determinada população. Representa a relação entre as pessoas com mais de 11 anos de estudo com a população acima dos 10 anos de idade.
DR	Doutores/cem mil habitantes. É um importante indicador da qualificação dos profissionais de ensino. Representa a relação entre o número de doutores em determinada região com a população em geral.
MSDR	Mestre e doutor/instituição de ensino superior. Corresponde ao número de mestres e doutores em determinada região geográfica, ponderado pelo número de instituições de ensino superior.
DOCENTESES	Docentes com ensino superior/total de professores educação básica (%). Corresponde ao percentual entre o número de professores com ensino superior atendendo à educação básica sobre o número total de professores.
ALPROF	Razão aluno/professor. Indica a relação entre o número de alunos pelo número de professores. É um importante indicador dos recursos educacionais.
ALTUR	Número de alunos/turma. Corresponde ao número de alunos matriculados sobre número de turmas das instituições.
INSTAL	Número de instituições/1.000 alunos. É obtido pela relação entre o número de instituições de ensino a cada mil alunos matriculados.
MATRPROF	Relação matrícula/professor ensino superior. Corresponde ao número de alunos matriculados no ensino superior para cada professor.
ALES	Alunos matriculados ensino superior/população (%). Indica a proporção da população que consegue atingir o ensino superior, apontado pela razão entre o número de matrículas no ensino superior e a população em geral.
GASTOS	Gastos do Fundo Nacional de Desenvolvimento da Educação (FNDE) com educação (mil reais/aluno matriculado). Esse indicador é obtido pela razão entre o montante despendido para a educação em relação ao somatório de alunos matriculados na educação básica e no ensino superior.

FONTE: Rigotti e Cerqueira (2004).
 Riani e Golgher (2004).
 Rigotti (2004).

4.2 Técnica de determinação dos escores fatoriais¹²

Como já abordado anteriormente, a mensuração do desenvolvimento educacional que vise representar, de forma mais fidedigna possível, a realidade de uma unidade geográfica deve envolver a análise de uma multiplicidade de variáveis, o que requer a utilização de uma ferramenta que possibilite o tratamento de um grande número de variáveis, reduzindo essas dimensões para um reduzido número de dimensões, mas sem perder informações relevantes do fenômeno em questão. Para tanto, utiliza-se a análise fatorial.

Antes de se entrar, propriamente, no mérito da análise fatorial, é necessário efetuar a análise de confiabilidade do constructo utilizado para a obtenção dos dados. A confiabilidade é caracterizada como o grau em que uma escala produz resultados consistentes entre medidas repetidas ou equivalentes de um mesmo objeto, relevando-se a ausência do erro (CORRAR; PAULO; DIAS FILHO, 2007). Trata-se da análise das escalas de mensuração, que permite determinar a extensão em que os itens estão relacionados com os demais e a fidedignidade do constructo. Portanto, conforme Corrar, Paulo e Dias Filho (2007), um modelo bastante utilizado é o Alfa de Cronbach (α), que trata da consistência interna baseada na correlação média entre os itens. Ainda, de acordo com Hair *et al.* (2006), a ideia principal da medida de consistência interna é que os itens ou indicadores individuais da escala devem medir o mesmo constructo e, assim, serem altamente intercorrelacionados.

Geometricamente, o Alfa de Cronbach pode ser calculado pela seguinte fórmula:

$$\alpha = \frac{k(\text{cov}/\text{var})}{1+(k-1)\left(\frac{\text{cov}}{\text{var}}\right)} \quad (1)$$

em que k é o número de variáveis consideradas; cov , a média das covariâncias; e var , a média das variâncias.

O valor assumido por α está entre 0 e 1, e, quanto mais próximo de 1, maior a fidedignidade das dimensões do constructo. No entanto, não há uma consistência quanto a um valor mínimo aceitável para essa estatística. Contudo, Hair *et al.* (2006) consideram 0,7 um mínimo ideal para pesquisas aplicadas, mas também podem ser aceitos valores próximos a 0,6 para pesquisas exploratórias.

¹² Para os cálculos dos modelos apresentados nesta seção, utiliza-se o pacote estatístico Stata (Statistics Data Analysis) 10.1.

O método da análise fatorial¹³ é uma técnica estatística multivariada que tem como objetivo a transformação de um número relativamente grande de variáveis em um número reduzido de fatores que possam explicar, de forma simples, as variáveis originais (MANLY, 1986). No ambiente econômico, essas variáveis estão, de alguma maneira, correlacionadas.

Além da redução do número de variáveis, a análise fatorial é uma técnica de interdependência, isto é, não há explicitada uma variável dependente, empregando, assim, o conceito de variável estática (KLEFENS, 2009). Porém, segundo Hair *et al.* (2006), esse conceito não é utilizado para prever uma variável dependente, mas, sim, para maximizar o poder de explicação do conjunto total de variáveis.

Desse modo, a análise fatorial pode ser realizada através do método de componentes principais¹⁴, que faz com que o primeiro fator contenha o maior percentual de explicação da variância total das variáveis da amostra, que o segundo fator contenha o segundo maior percentual, e assim sucessivamente (FERREIRA JUNIOR; BAPTISTA; LIMA, 2004).

No modelo apresentado a seguir, cada variável observada X_i , representa uma combinação linear dos n componentes principais. Desse modo, cada uma das i -ésimas variáveis é uma combinação de k ($k < n$) fatores comuns e de um fator específico. Genericamente, um modelo de análise fatorial pode ser apresentado da seguinte forma, considerando que cada variável se relaciona, linearmente, com k fatores comuns (F), um fator único (U) e um termo aleatório:

$$X_i = A_{i1}F_1 + A_{i2}F_2 + \dots + A_{ik}F_k + U_i + E_i \quad (2)$$

em que X_i são as variáveis originais; A_{ik} , as cargas fatoriais, usadas para combinar linearmente os fatores comuns; F_1, F_2, \dots, F_k , os fatores comuns; U_i , o fator único; e E_i , o termo de erro que capta a variação específica de X_i não explicada pela combinação linear das cargas fatoriais com os fatores comuns.

As cargas fatoriais A_{ik} indicam a intensidade das relações entre as variáveis normalizadas X_i e os fatores. Quanto maior uma carga fatorial, mais associada com o fator se encontra a variável. O quadrado das cargas fatoriais representa a contribuição relativa de cada fator para a variância total da

¹³ A técnica da análise fatorial surgiu com Charles Spearman, em 1904, para descrever a inteligência através de um único fator. Spearman (1904) desenvolveu um método para a criação de um índice geral de inteligência (fator "g"), com base nos resultados de vários testes que avaliavam essa aptidão. No entanto, o termo análise fatorial foi introduzido por Thurstone (1931).

¹⁴ Segundo Corrar, Paulo e Dias Filho (2007), o método da Análise de Componentes Principais procura uma combinação linear entre as variáveis, de forma que o máximo da variância seja explicada.

variável. Por outro lado, para cada variável, a soma dessas cargas fatoriais ao quadrado oferece a estimativa da comunalidade, que indica a proporção da variância total de cada variável que é explicada pelo conjunto de fatores comuns. O Eigenvalue, ou raiz característica, expressa a variância total do modelo explicada por cada fator. Seu valor é dado pelo somatório dos quadrados das cargas fatoriais de cada variável associada ao fator específico. A raiz característica dividida pelo número de variáveis (X_i) determina a proporção da variância total explicada pelo fator.

Para determinar o número de fatores a serem utilizados pela matriz de correlação R , extraem-se apenas os fatores associados a raízes características maiores do que a unidade (ZAMBRANO; LIMA, 2004). Segundo Hair *et al.* (2006), com a adoção de um número muito reduzido de fatores, não é possível identificar estruturas importantes existentes nos dados, e, por outro lado, se o número é excessivo, pode vir a ter problemas de interpretabilidade dos dados.

Estimadas as cargas fatoriais, com o intuito de obter uma estrutura de fatores mais simplificada, o próximo passo consiste em efetuar a rotação dos fatores. De acordo com Hair *et al.* (2006), o efeito final de rotacionar a matriz fatorial é redistribuir a variância dos primeiros fatores para os últimos, com a intenção de atingir um padrão fatorial mais simples e mais significativo. Essa rotação altera a parcela de contribuição de cada fator, mas mantém os valores das comunalidades e a proporção de variância explicada pelo conjunto de fatores, dado que ela maximiza a carga de uma variável em um único fator, permitindo, assim, uma melhor representatividade de cada fator.

Os métodos de rotação fatorial ortogonal¹⁵ que se destacam na literatura e são encontrados em Mingoti (2005) são: Critério Varimax, Quatimax e Orthomax¹⁶. Conforme Hoffmann (1999), a técnica Varimax é um dos critérios mais usados nos estudos de análise fatorial envolvendo a transformação ortogonal e, segundo Mingoti (2005), produz soluções mais práticas que as outras técnicas. Esse método de rotação foi o utilizado no presente estudo. Após calcular as cargas fatoriais e identificar os fatores, estimam-se os escores fatoriais — variável importante para a elaboração do IDE.

Desse modo, o escore para cada observação (UF) é o resultado da multiplicação do valor padronizado das variáveis, pelo coeficiente do escore

¹⁵ Rotação fatorial ortogonal é a rotação fatorial, computada de forma que os fatores sejam extraídos de modo que seus eixos se mantenham em 90 graus. Assim, cada fator se mantém independente, ou ortogonal, em relação aos demais (ANDERSON *et al.*, 2009).

¹⁶ Para o desenvolvimento algébrico dos critérios de rotação dos fatores, ver Mingoti (2005, p. 121-130).

fatorial correspondente, sendo a expressão geral para estimação do j -ésimo fator F_j dada por:

$$F_j = W_{j1}X_1 + W_{j2}X_2 + W_{j3}X_3 + \dots + W_{jn}X_n \quad (3)$$

em que os W_{ji} são os coeficientes dos escores fatoriais, e n , o número de variáveis.

Monteiro e Pinheiro (2004) apontam que, quanto mais distante de zero for o escore fatorial de uma observação, em valores positivos, melhor será a posição relativa da observação em um fator. Portanto, o desempenho dos estados em cada fator é determinado pelos valores comparativos da coluna correspondente ao fator na matriz de escores fatoriais.

Por fim, cabe realizar a análise da qualidade do ajuste do modelo de análise fatorial. Para tanto, foram utilizadas duas medidas para examinar as intercorrelações entre as variáveis. O primeiro é o Teste de Esfericidade de Bartlett (Bartlett Test of Sphericity), o qual testa a hipótese de que as variáveis sejam não correlacionadas na população. A hipótese básica afirma que a matriz de correlação da população é uma matriz identidade, o que indica que o modelo fatorial é inapropriado. Para prosseguir com a análise, é preciso que a hipótese de que não há correlação entre as variáveis seja rejeitada¹⁷.

O outro indicador verifica o grau de intercorrelações entre as variáveis e a adequação da AF. A estatística de Kaiser-Meyer-Olkin (KMO) é um indicador que compara a magnitude do coeficiente de correlação observado com a magnitude do coeficiente de correlação parcial¹⁸. Considerando que os valores deste variam de 0 a 1, valores para KMO menores do que 0,5 indicam a não adequabilidade da análise. Valores entre 0,5 e 0,6 indicam uma adequabilidade ruim do modelo; entre 0,6 e 0,7, regular; entre 0,7 e 0,8, ruim; entre 0,8 e 0,9, ótimo; e entre 0,9 e 1,0, excelente (HAIR *et al.*, 2006).

Desse modo, a elaboração do Índice de Desenvolvimento Educacional de cada estado será feita através dos escores fatoriais, ou seja, dos valores dos fatores de cada uma das 27 observações (estados). O método para a elaboração do IDE será desenvolvido na subseção seguinte.

¹⁷ A estatística desse teste é dada por: $\chi^2 = - \left[(n-1) - \frac{2p+5}{6} \right] \ln |R|$, que tem uma distribuição qui-quadrado com $v = \frac{p(p-1)}{2}$ graus de liberdade, em que n é o tamanho da amostra; p , o número de variáveis; e $|R|$, o determinante da matriz de correlação.

¹⁸ A medida de KMO é dada por: $KMO = \frac{\sum_{i \neq j} R_{ij}^2}{\sum_{i \neq j} R_{ij}^2 + \sum_{i \neq j} Q_{ij}^2}$, em que R_{ij} é a correlação amostral entre as variáveis X_i e X_j ; e Q_{ij} , a correlação parcial entre X_i e X_j .

4.3 Método de cálculo do Índice de Desenvolvimento Educacional

O IDE é calculado a partir da soma dos escores fatoriais, multiplicados pelas raízes características (autovalores) dos fatores comuns ortogonais. No entanto, espera-se que os escores associados aos estados tenham distribuição simétrica, em torno da média zero. Desse modo, metade apresenta sinais negativos; e a outra metade, sinais positivos, sendo que os de menor desenvolvimento educacional apresentarão valores negativos.

Uma vez obtidos os escores fatoriais, a proporção da variância explicada pelos fatores extraídos e os F_{ij} , é possível realizar o cálculo do Índice Bruto de Desenvolvimento Educacional (IBDE), dado pela seguinte equação:

$$IBDE_i = \sum_{j=1}^p \frac{\sigma_j^2}{\sum \sigma_j^2} F_{ij}, \quad (4)$$

em que $IBDE_i$ é o Índice Bruto de Desenvolvimento Educacional do i -ésimo estado; σ_j^2 , a variância explicada pelo j -ésimo fator; p , o número de fatores utilizados na análise; e F_{ij} , o j -ésimo escore fatorial do i -ésimo estado.

Por fim, obtido o $IBDE_i$, é possível calcular o Índice de Desenvolvimento Educacional, utilizando-se a seguinte equação:

$$IDE_i = \frac{(IBDE_i - IBDE^{min})}{(IBDE^{max} - IBDE^{min})} \quad (5)$$

em que $IBDE_i$ é o valor do Índice Bruto de Desenvolvimento Educacional do i -ésimo estado analisado, e $IBDE^{min}$ e $IBDE^{max}$ são os valores mínimos e máximos, respectivamente, obtidos para o $IBDE_i$ do i -ésimo estado.

Desse modo, o IDE_i constitui-se de um índice que tem como referência o melhor valor de $IBDE_i$ dentre os 27 estados analisados. Assim, o estado com melhor desempenho da educação apresentará IDE_i igual a 1 (um), e os demais apresentarão resultados inferiores, conforme o seu desenvolvimento, para cada período analisado¹⁹.

E, em um último tópico de análise, foram consideradas três categorias de grau de desenvolvimento educacional. Aqueles estados cujos resultados se situam meio desvio-padrão acima da média serão rotulados como de alto grau de desenvolvimento educacional; como desenvolvimento médio, serão classificados os situados entre meio desvio-padrão acima e abaixo da média; e aqueles com valores inferiores a meio desvio-padrão abaixo da média serão classificados como de baixo desenvolvimento educacional.

¹⁹ Como se trata de um índice que varia entre 0 e 1, o estado que apresentar o pior índice será igual a zero, enquanto, no outro extremo, o melhor será igual a 1.

5 Resultados e discussões

Nesta seção, estão os resultados da aplicação da análise fatorial na determinação do IDE, para o período 2007-11.

A primeira estatística a ser analisada é o Alfa de Cronbach, o qual testa a consistência interna dos dados. Esses resultados estão apresentados na Tabela 1.

Tabela 1

Estatística Alfa de Cronbach — 2007-11

DISCRIMINAÇÃO	2007	2008	2009	2010	2011
Estatística α	0,8683	0,8598	0,8442	0,8431	0,8396

FONTE DOS DADOS BRUTOS: Censo de Educação do INEP (2007-2011).
IPEA.

Sistema de Informações sobre Orçamentos Públicos em Educação.

NOTA: Valores padronizados para a estatística Alfa de Cronbach. Em adição, também foram analisadas as estatísticas Alfa de Cronbach para as dimensões apresentadas no Quadro 1, que, por restrições de espaço, foram omitidas. No entanto, suas estatísticas encontram-se entre 0,50 e 0,80, em todos os casos (com exceção da dimensão capital humano referente a 2011, com estatística de 0,4662).

Através das estatísticas de α presentes na Tabela 1, para o período 2007-11, infere-se que as intercorrelações entre os itens apresentam valores acima de 0,7 e, por isso, podem ser consideradas um bom ajuste, dada a dificuldade em se conseguir consolidar uma base de dados com essas variáveis.

A aplicação da análise fatorial pelo método dos componentes principais à base de dados, para o período em questão, possibilitou a extração de cinco fatores com raiz característica maior que a unidade e que sintetizam as informações contidas nas 17 variáveis originais para o ano de 2007, sendo, do mesmo modo, para os anos de 2008 e 2010, conforme o Quadro 2. No entanto, para o ano de 2009, extraíram-se quatro fatores com raiz característica maior que a unidade, e, para o ano de 2011, extraíram-se seis fatores com essa característica²⁰.

Antes de analisar a participação de cada fator na explicação da variância total, cabe verificar a qualidade dos resultados da aplicação da análise fatorial. Para determinar se os dados suportam essa aplicação, é aplicado o Teste de Esfericidade de Bartlett, cujo objetivo é constatar a presença de correlações entre as variáveis. Através dos resultados do Quadro 2, para

²⁰ Na literatura sobre análise multivariada, não se encontram informações sobre uma regra que deve ser seguida quanto à determinação do número de fatores. O critério adotado para a presente pesquisa é o requisito que a raiz característica associada ao fator seja maior que a unidade.

esse teste, são encontrados valores elevados, ou seja, 514,02 para 2007; 458,11 para 2008; 436,42 para 2009; 455,04 para 2010; e 456,67 para 2011. Assim, sob o nível de significância de 1%, a matriz de correlação entre as variáveis difere de uma matriz identidade²¹.

Quadro 2

Raiz característica, percentual explicado por cada fator e variância acumulada — 2007-11

ANOS	FATORES	RAIZ CARACTERÍSTICA	VARIÂNCIA EXPLICADA PELO FATOR (%)	VARIÂNCIA ACUMULADA (%)	TESTE DE ESFERICIDADE DE BARTLETT ($\chi^2(136)$)	ESTATÍSTICA KMO
2007	F1	6,7763	39,86	39,86	514,02	0,5137
	F2	3,2809	19,30	59,16		
	F3	1,8907	11,12	70,28		
	F4	1,3158	7,74	78,02		
	F5	1,1391	6,70	84,72		
2008	F1	6,4187	37,76	37,76	458,11	0,6566
	F2	3,4046	20,03	57,78		
	F3	1,7748	10,44	68,22		
	F4	1,4988	8,82	77,04		
	F5	1,1482	6,75	83,79		
2009	F1	6,3492	37,35	37,35	436,42	0,6503
	F2	3,4019	20,01	57,36		
	F3	1,9893	11,70	69,06		
	F4	1,4493	8,53	77,59		
2010	F1	6,5360	38,45	38,45	455,04	0,5976
	F2	2,9701	17,47	55,92		
	F3	2,0133	11,84	67,76		
	F4	1,4122	8,31	76,07		
	F5	1,1396	6,70	82,77		
2011	F1	6,3033	37,08	37,08	456,67	0,4903
	F2	2,8671	16,87	53,94		
	F3	2,1165	12,45	66,39		
	F4	1,3932	8,20	74,59		
	F5	1,1630	6,84	81,43		
	F6	1,0401	6,12	87,55		

FONTE DOS DADOS BRUTOS: Censo de Educação do INEP (2007-2011).

IPEA.

Sistema de Informações sobre Orçamentos Públicos em Educação.

²¹ Cada um desses valores se refere à estatística chi-quadrado, com 136 graus de liberdade. Ao nível de significância de 1%, a estatística crítica de chi-quadrado com 136 graus de liberdade é de 135,807, que faz com que o valor calculado seja maior que o crítico, não se aceitando, assim, a hipótese nula da matriz de correlação ser igual a uma matriz identidade.

Em seguida, como medida de adequabilidade amostral, empregou-se a estatística KMO. Os resultados calculados são de 0,5137, 0,6566, 0,6503, 0,5976 e 0,4903 para os anos de 2007 a 2011 respectivamente. Desse modo, baseando-se na classificação fornecida por Hair *et al.* (2006), estatísticas acima de 0,50 indicam que os dados são adequados à realização da análise fatorial. Para o presente caso, para o período 2007-10, os valores são superiores a 0,5, o que indica que os dados são adequados à realização da análise fatorial. Quanto à estatística KMO para 2011, como seu valor se encontra muito próximo a 0,5, não se impossibilita essa aplicação²². Então, perante esses testes, conclui-se que as amostras de dados são passíveis de serem analisadas com a aplicação da análise fatorial.

Ainda conforme o Quadro 2, percebe-se uma forte concentração da participação do Fator 1 na explicação da variância total. Esses resultados mostram que, para o ano de 2007, os cinco fatores com raiz característica maior que um explicam, conjuntamente, 84,72% da variância total; para o ano de 2008, esses mesmos cinco fatores representam 83,79% da variância. No entanto, no ano de 2009, foram identificados quatro fatores com raiz característica maior que um, os quais representam 77,59% da variância total. Já em 2010, foram, novamente, identificados cinco fatores, que expressam 82,77% da variância total, e, por fim, para 2011, foram identificados seis fatores, os quais explicam 87,55% da variância total das variáveis.

As matrizes com as cargas fatoriais para os anos de 2007 e 2011 encontram-se, respectivamente, nas Tabelas 2 e 3, nas quais estão identificadas as variáveis que compõem cada um dos fatores, o grau de correlação entre cada variável e cada fator, bem como a comunalidade.²³

Para a interpretação das cargas fatoriais, foram consideradas apenas aquelas com valores superiores a 0,6, as quais se encontram destacadas nas tabelas. Os valores encontrados para as comunalidades revelam que, praticamente, todas as variáveis têm sua variabilidade captada e representada pelos fatores apontados. Desse modo, tem-se que a menor comunalidade encontrada para 2007 é de 60,48%, enquanto, para o ano de 2011, é de 69,58%.

A partir dos dados da Tabela 2, observa-se que o Fator 1 possui a maior parcela de variância entre os fatores obtidos e se encontra, forte e negativamente, correlacionado com as variáveis TANALF, DISTEF e DISTEM e correlaciona-se, com menos intensidade, com o número de instituições de ensino a cada mil alunos. Esses três primeiros indicadores estão

²² Os demais testes são favoráveis à aplicação da análise fatorial.

²³ Por razões de espaço, as cargas fatoriais para os anos de 2008, 2009 e 2010 não foram apresentadas.

estritamente relacionados ao desempenho dos alunos; então, essa forte relação negativa irá dar um peso mais elevado aos estados que apresentam, por exemplo, maiores taxas de analfabetismo, induzindo a região a que pertence a uma situação comparativa inferior à dos demais estados. Por outro lado, o Fator 1 encontra-se positivamente correlacionado com as variáveis ANES, DR, DOCENTESES, ALES e GASTOS, as quais estão estritamente relacionadas à categoria capital humano, através da qual se infere que uma população com mais anos de estudo e com profissionais mais qualificados nas instituições faz a diferença no desenvolvimento regional. Além disso, a percentagem de alunos com acesso ao ensino superior e os gastos federais com educação são importantes fatores que contribuem para evidenciar essas diferenças, mesmo que esses gastos, em âmbito nacional, não estejam produzindo os resultados esperados.

Tabela 2

Cargas fatorias e comunalidades — 2007

VARIÁVEIS	CARGAS FATORIAIS					COMUNALIDADE
	F1	F2	F3	F4	F5	
TANALF	-0,8919	-0,1540	-0,0443	-0,1023	0,1266	0,8476
DISTEF	-0,8959	0,1991	0,1592	-0,0699	0,2426	0,9313
DISTEM	-0,8640	0,2604	0,1266	-0,0521	0,2871	0,9154
TESCBEF	-0,4801	0,0113	-0,5278	-0,0463	0,4754	0,7374
PROFEB	-0,2252	-0,8523	0,3713	-0,0200	0,1565	0,9399
ALUPROF	-0,4300	0,7610	-0,0271	0,2888	-0,1714	0,8776
ANES	0,8958	0,1701	0,2561	0,0093	0,1642	0,9240
ALTUR	-0,2006	0,8005	0,1833	0,3151	0,1834	0,8476
INSTAL	-0,6624	-0,5390	0,0862	-0,2477	-0,1066	0,8094
TRENDESC	0,4222	-0,5385	-0,2597	0,2704	-0,3350	0,7210
MATRPROF	-0,0588	0,3305	0,5600	0,3260	-0,2687	0,6048
DR	0,7992	0,2830	-0,0672	-0,4254	0,2052	0,9463
COMA	0,2411	-0,1132	-0,6476	0,4644	0,3194	0,8081
MSDR	0,1043	0,5595	-0,1227	-0,6860	-0,1879	0,8449
DOCENTESES	0,7449	0,2358	-0,4240	-0,0428	-0,0509	0,7947
ALES	0,8876	-0,0323	0,2698	-0,0092	0,2620	0,9303
GASTOS	0,6725	-0,1480	0,5011	0,0505	0,4414	0,9226

FONTE DOS DADOS BRUTOS: Censo de Educação do INEP (2007-2011).

IPEA.

Sistema de Informações sobre Orçamentos Públicos em Educação.

A participação dos demais fatores é de menor importância na explicação total, contudo cabe identificar as variáveis associadas a cada fator. Ao Fator 2, estão relacionadas as variáveis PROFEB, ALPROF e ALTUR, sendo estas associadas ao capital humano e a questões estruturais. Já o Fator 3 está relacionado somente com a variável COMA, referente ao percentual

de concluintes do ensino superior, e o Fator 4 está relacionado com o número de mestres e doutores por instituição de ensino superior (MSDR).

Na Tabela 3, estão presentes as cargas fatoriais para 2011. Nessa análise, de modo semelhante a 2007, verifica-se uma associação forte e negativa do Fator 1 (que expressa 37,08% da variância total) com as variáveis TANALF, DISTEF e DISTEM e uma correlação negativa, não tão intensa, com o número de instituições a cada mil alunos. Por outro lado, há um aumento da carga fatorial associado à participação na população das pessoas com mais de 11 anos de estudo e também existe uma relação positiva desse fator com as variáveis DOUTOR, DOCENTESES, ALES e GASTOS, mesma relação apresentada no período já analisado.

A relação do Fator 2, que explica 16,87% da variância total, ocorre, de forma negativa, com a variável PROFEB e, positiva, com as variáveis ALPROF e ALTUR. Ainda, o Fator 3 está relacionado, positivamente, com a variável MATRPROF, ou seja, associado a questões estruturais. Por fim, os Fatores 4, 5 e 6, encontram-se relacionados positivamente com as variáveis MSDR, TRENDESC e TESCBEF respectivamente. Essas três variáveis estão associadas, respectivamente, a questões de qualificação profissional e ao desempenho dos alunos nos diversos níveis de ensino.

Tabela 3

Cargas fatoriais e comunalidades — 2011

VARIÁVEIS	CARGAS FATORIAIS						COMUNALIDADE
	F1	F2	F3	F4	F5	F6	
TANALF	-0,8643	-0,1833	0,1719	-0,0148	0,0651	-0,0366	0,8160
DISTEF	-0,8443	0,1837	0,2597	0,2874	0,0238	0,0873	0,9048
DISTEM	-0,8574	0,2527	0,2019	0,2183	-0,1042	-0,0201	0,8986
TESCBEF	0,3071	-0,2312	0,3108	0,1726	0,2263	0,7180	0,8409
PROFEB	-0,2176	-0,8346	0,3614	0,1160	-0,1968	0,1082	0,9384
ALPROF	-0,4018	0,7125	0,3059	-0,3151	0,1712	0,1542	0,9150
ANES	0,9013	0,0566	0,2310	0,1144	0,0282	-0,0564	0,8859
ALTUR	-0,0265	0,7589	0,4699	-0,3087	-0,0875	-0,0181	0,9007
INSTAL	-0,6651	-0,5939	0,0546	0,1923	0,2895	-0,1925	0,9560
TRENDESC	0,2356	-0,3821	0,0036	-0,4277	0,6876	0,0630	0,8612
MATRPROF	0,0042	0,0475	0,8595	0,0187	0,1975	-0,0417	0,7821
DR	0,7567	0,2386	-0,1456	0,4501	0,1150	-0,1950	0,9046
COMA	0,3422	-0,2164	0,2960	-0,3374	0,1878	-0,5433	0,6958
MSDR	-0,0942	0,4987	-0,1257	0,6138	0,5449	-0,1341	0,9650
DOCENTESES ...	0,7913	0,1645	-0,2890	-0,1228	0,0870	0,2697	0,8321
ALES	0,7959	-0,1285	0,4377	0,2349	-0,0378	-0,0357	0,8995
GASTOS	0,7321	-0,0470	0,4923	0,1677	-0,2639	-0,0912	0,8867

FONTE DOS DADOS BRUTOS: Censo de Educação do INEP (2007-2011).

IPEA.

Sistema de Informações sobre Orçamentos Públicos em Educação.

Nessa perspectiva, além da utilização da AF para a extração das cargas fatoriais e dos autovalores, sua metodologia também tem como intuito identificar as variáveis que compõem cada um dos fatores, o que possibilita eleger um rótulo para cada fator. A redução da dimensionalidade e a constituição dos fatores para o período em análise são apresentadas no Quadro 3.

Desse modo, os fatores comuns aos períodos analisados são rotulados da seguinte forma²⁴: o Fator 1 é identificado como inerente a características desempenho-estruturais, sendo que, ao longo do período, representa uma variação total média 38,1% do modelo. Existe grande concentração das variáveis abordadas nesse fator, o que justifica sua proporção na explicação total da variância. As correlações dos indicadores associadas a esse fator são elevadas, próximas a 0,90, sendo as principais a taxa de analfabetismo e o grau de distorção (correlações negativas) e anos de estudo, alunos no ensino superior e o número de doutores (correlações positivas).

Quadro 3

Descrição dos fatores

ANOS	INDICADORES			
	Fator 1	Fator 2	Fator 3	Fator 4
2007	tanalf, distef, distem, anes, instal, dr, docenteses, ales, gastos	profef, aluprof, altur	matrprof, coma	msdr
2008	tanalf, distef, distem, anes, instal, dr, docenteses, ales, gastos	profef, aluprof, altur	matrprof	trendesc
2009	tanalf, distef, distem, anes, instal, dr, docenteses, ales, gastos	profef, aluprof, altur	matrprof	msdr
2010	tanalf, distef, distem, anes, instal, dr, docenteses, ales, gastos	profef, aluprof, altur, msdr	matrprof	trendesc
2011	tanalf, distef, distem, anes, instal, dr, docenteses, ales, gastos	profef, alprof, altur	matrprof	msdr

Já o Fator 2 é responsável pela alocação estrutural do ensino básico, o qual explica, em média, 18,74% da variação total do modelo. As variáveis que o compõem, a saber, razão aluno/professor, alunos por turma e o número de professores da educação básica, estão conectadas à questão estrutural, problema este enfrentado por muitas escolas brasileiras.

²⁴ O processo de “rotular” os fatores concentrar-se-á nos quatro primeiros fatores, por dois motivos principais: primeiramente, porque, ao longo do período de análise, tais fatores explicam, em média, 76,66% da variância; em segundo lugar, pelo fato de tais fatores apresentarem alternância relativamente baixa das variáveis que os compõem. Além disso, outro aspecto relevante de ser mencionado, o qual é uma limitação do presente estudo, dada a quantidade relativamente pequena de indicadores, é que alguns fatores foram representados por apenas uma variável.

Por fim, o Fator 3 está associado ao número de matrículas por professor no ensino superior, correspondendo a 11,51% da variação total do modelo, e o Fator 4 representa, em média, 8,32% da variação total do modelo, podendo ser identificado como associado ao grau de qualificação dos profissionais do ensino superior. Dada a composição desses fatores, o melhor rótulo para tais é o próprio indicador que os compõe.

Com isso, pode-se observar a importância e o relacionamento individual de cada indicador com os fatores que formarão o IDE. Uma vez verificadas as cargas fatoriais, o passo seguinte é observar os escores fatoriais, ou seja, o valor do fator para cada estado.

No Quadro 4, estão representados o Índice brasileiro de desenvolvimento educacional (IBDE), resultante da aplicação da Equação 4 aos escores fatoriais, o IDE, resultante da Equação 5, e o *ranking* dos estados brasileiros. A interpretação do IDE segue a lógica de que os escores próximos a um (unidade) indicam grau de desenvolvimento elevado, por outro lado, quanto mais próximo a zero, mais atrasado é o estado em comparação aos demais.

O IDE nacional médio situou-se em 0,433, ao longo do período. Desse modo, tanto através do Quadro 4, quanto através da Figura 5, percebe-se o desenho de um mapa geográfico que coloca os estados da Região Nordeste no extremo inferior, e, por outro lado, o extremo superior apresenta, principalmente, o Distrito Federal e os Estados de São Paulo e Rio de Janeiro como os mais desenvolvidos, quando o quesito analisado é a educação. O fato de o Distrito Federal ocupar o topo desse *ranking* pode ser explicado pelo comportamento de seus próprios indicadores, se analisados individualmente. Conforme visto anteriormente, é o que apresenta a menor taxa de analfabetismo; tem uma das melhores estatísticas de distorção idade-série; apresenta a maior porcentagem da população com mais de 11 anos de estudo, quando analisadas estatísticas associadas a questões estruturais e — principalmente — de capital humano; além de possuir a maior concentração de professores com grau de qualificação elevado ponderado pela população. Voltando ao lado econômico, é a região geográfica que detém a maior relação entre gastos federais por aluno matriculado em todos os níveis de ensino.

Quadro 4

Índice brasileiro de desenvolvimento educacional (IBDE), Índice de Desenvolvimento Educacional (IDE) e *ranking* (R) dos estados brasileiros e do Brasil — 2007-11

UF	2007			2008			2009			2010			2011		
	IBDE	IDE	R												
DF	1,752	1,000	1	1,482	1,000	1	1,760	1,000	1	1,772	1,000	1	1,517	1,000	1
SP	0,940	0,709	2	0,882	0,747	3	0,841	0,655	3	0,818	0,650	3	0,715	0,693	2
RJ	0,660	0,609	3	0,927	0,766	2	1,017	0,721	2	0,855	0,664	2	0,518	0,618	3
MS	0,240	0,458	7	0,416	0,550	6	0,331	0,463	5	0,217	0,430	6	0,481	0,604	4
AP	0,130	0,419	10	0,455	0,566	4	0,269	0,440	7	0,158	0,408	8	0,411	0,577	5
RS	0,189	0,440	8	0,327	0,512	7	0,637	0,578	4	0,474	0,524	4	0,269	0,523	6
RR	-0,150	0,318	18	-0,435	0,190	22	-0,176	0,273	18	-0,052	0,331	16	0,269	0,523	7
PR	0,427	0,525	4	0,420	0,551	5	0,322	0,460	6	0,321	0,468	5	0,189	0,492	8
SE	-0,121	0,329	16	-0,010	0,370	15	0,066	0,363	12	-0,041	0,335	15	0,185	0,491	9
SC	0,094	0,406	13	0,165	0,444	11	0,210	0,417	8	0,132	0,399	10	0,145	0,475	10
ES	0,107	0,410	12	0,035	0,389	14	-0,081	0,308	16	0,038	0,364	13	0,083	0,452	11
RN	-0,236	0,287	19	-0,177	0,299	17	0,155	0,397	10	0,046	0,367	12	0,045	0,437	13
GO	0,277	0,471	6	0,173	0,447	10	0,069	0,365	11	0,141	0,402	9	0,043	0,436	14
RO	0,054	0,391	14	-0,140	0,315	16	-0,139	0,287	17	-0,223	0,268	18	-0,012	0,415	15
MT	-0,002	0,371	15	0,133	0,430	13	-0,015	0,333	14	0,032	0,362	14	-0,023	0,411	16
MG	0,121	0,416	11	0,174	0,447	9	0,035	0,352	13	0,058	0,371	11	-0,111	0,377	17
PB	-0,737	0,108	26	-0,535	0,148	25	-0,419	0,181	21	-0,483	0,173	25	-0,135	0,368	18
CE	-0,378	0,237	22	-0,464	0,178	23	-0,409	0,185	20	-0,326	0,231	23	-0,163	0,358	19
TO	-0,409	0,225	23	-0,333	0,233	19	-0,442	0,173	24	-0,602	0,129	26	-0,205	0,341	20
AM	0,316	0,485	5	0,156	0,440	12	-0,062	0,315	15	-0,088	0,318	17	-0,217	0,337	21
PE	-0,139	0,322	17	-0,279	0,256	18	-0,346	0,209	19	-0,267	0,252	19	-0,261	0,320	22
AC	-0,255	0,281	20	-0,478	0,172	24	-0,514	0,145	25	-0,278	0,248	20	-0,278	0,314	23
PA	-0,279	0,272	21	-0,418	0,197	20	-0,441	0,173	23	-0,278	0,248	21	-0,301	0,305	24
AL	-0,410	0,225	24	-0,428	0,193	21	-0,429	0,177	22	-0,304	0,238	22	-0,511	0,224	25
BA	-0,485	0,198	25	-0,587	0,126	26	-0,614	0,108	26	-0,426	0,194	24	-0,750	0,133	26
MA	-0,847	0,069	27	-0,778	0,045	27	-0,897	0,001	27	-0,955	0,000	28	-0,882	0,083	27
PI	-1,039	0,000	28	-0,885	0,000	28	-0,901	0,000	28	-0,916	0,014	27	-1,098	0,000	28
BR	0,182	0,437	9	0,202	0,459	8	0,176	0,405	9	0,176	0,415	7	0,074	0,448	12

FONTE DOS DADOS BRUTOS: Censo de Educação do INEP (2007-2011).

IPEA.

Sistema de Informações sobre Orçamentos Públicos em Educação.

Em contrapartida, a concentração dos estados da Região Nordeste na parte inferior do Quadro 4 é explicada por suas altas taxas de analfabetismo, em torno de 19%, em 2011, pela acentuada diferença no grau de qualificação dos professores (por exemplo, na Região Nordeste, a proporção de professores doutores a cada 100 mil habitantes é de 38,15; enquanto, na Sudeste, essa proporção é de 63,92; já na Região Sul, é de 69,80) e pela baixa parcela da população com 11 anos ou mais de estudo (na Região Nordeste, essa estatística é de 26,50%, enquanto, nas Regiões Sudeste, Sul e Centro-Oeste, é de 38,50%, 36,15% e 38,96% respectivamente).

Em uma análise individual de alguns estados, percebe-se que o Distrito Federal ocupa o topo do *ranking* ao longo de todo o período, seguido por São Paulo e Rio de Janeiro. Alguns estados em particular merecem destaque, pela melhora em sua colocação ao longo do período. Amapá e Rio Grande do Norte ocupavam, respectivamente, a 10.^a e a 19.^a posição em 2007, já, em 2011, ocuparam a 5.^a e a 13.^a posição respectivamente. Essa melhora no desempenho é fruto da evolução em indicadores, como taxa de analfabetismo, proporção da população com mais de 11 anos de estudo, melhor qualificação dos docentes, dentre outros. Por outro lado, alguns estados apresentaram uma brusca queda no *ranking*, de que são exemplos Goiás, Minas Gerais e Amazonas.

A parte inferior do *ranking* permanece ocupada pelo Piauí no período, seguido por Maranhão, Bahia e Alagoas. Particularmente, esses quatro estados praticamente não deixaram as quatro últimas colocações desde 2007. Esse desempenho é reflexo das altas taxas de analfabetismo apresentadas, não apenas por estes, mas por toda a Região Nordeste, por apresentar baixos investimentos em educação²⁵ e também pela baixa qualificação dos profissionais que atuam nas instituições. De modo geral, os demais estados não apresentaram grandes oscilações em sua posição, mantendo-se estáveis ao longo da série.

Para uma análise mais agregada, a Figura 5 apresenta o IDE médio para o período 2007-11 para os estados brasileiros, o qual se situou em 0,371, com um desvio-padrão de 0,2025, e, a partir daí, foram elaboradas categorias de desenvolvimento da educação. Assim, os estados que apresentaram um IDE superior à média mais meio desvio-padrão (0,472) foram classificados como de alto grau de desenvolvimento. Já os estados que se enquadraram no intervalo de 0,471 a 0,280 foram classificados como de

²⁵ O baixo grau de investimento federal em determinada região, neste caso, a Nordeste, afeta também o desempenho das demais regiões. Isso ocorre porque há uma transferência intrarregional de recursos, e como não há uma base adequada para gerar resultados a esses investimentos nessa região, ocorre um comprometimento do desenvolvimento das demais regiões, em função dessa má alocação de recursos.

médio desenvolvimento educacional, e, por fim, os que apresentaram IDE abaixo de 0,270 foram classificados como de baixo desenvolvimento educacional.

O total de estados pertencentes à categoria de alto desenvolvimento chega a apenas sete (Distrito Federal, São Paulo, Rio de Janeiro, Rio Grande do Sul, Mato Grosso do Sul, Paraná e Amapá)²⁶, o que significa que apenas um seleto grupo apresenta um desenvolvimento educacional destacável no País. Nessa categoria, não se percebe a presença de estados da Região Nordeste.

Os estados que apresentam IDE médio são 11 (Santa Catarina, Goiás, Minas Gerais, Espírito Santo, Mato Grosso, Amazonas, Sergipe, Rio Grande do Norte, Rondônia, Roraima e Pernambuco), sendo a grande parte das UF classificadas nessa categoria. E novamente, a Região Nordeste, juntamente com alguns estados da Região Norte, predomina na categoria de baixo IDE. Nessa categoria, estão presentes nove estados (Pará, Ceará, Acre, Tocantins, Alagoas, Paraíba, Bahia, Maranhão e Piauí).

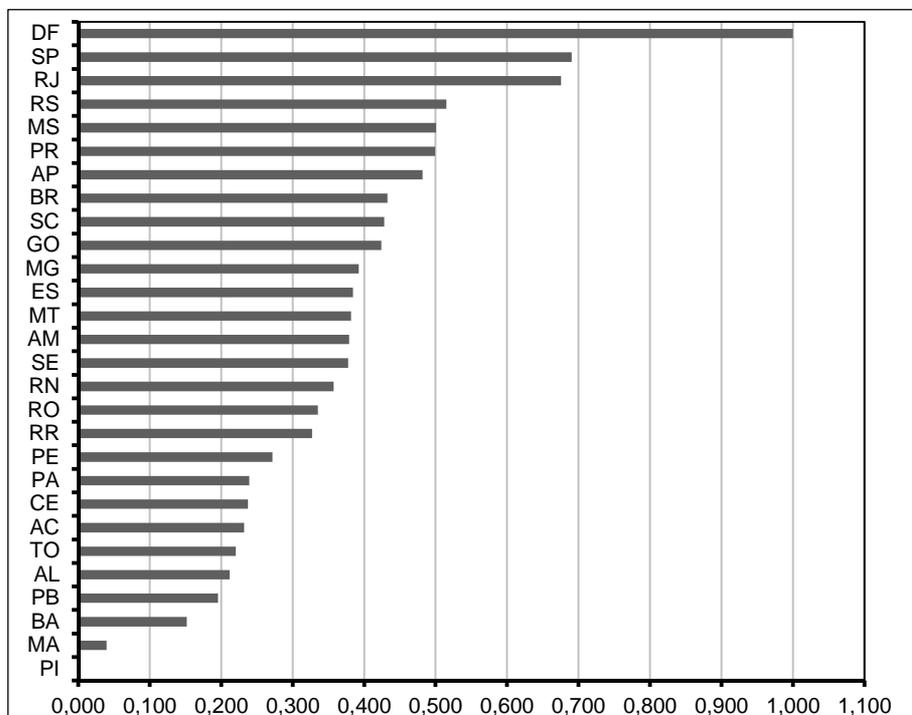
Assim, os principais resultados desta seção confirmam as evidências levantadas nas primeiras seções, apontando que o desempenho do sistema educacional brasileiro necessita de ajustes urgentes. E assim, a educação pode ser tomada como um ponto de partida para a redução das desigualdades de salário e renda existentes entre as regiões brasileiras.

Nesse aspecto, os resultados apresentados nesta seção vêm evidenciar, ainda mais, as discrepâncias estaduais e regionais apontadas anteriormente e o baixo dinamismo para reverter tal situação, particularmente pela Região Nordeste, de modo que, através de um índice ou de uma classificação, se torna mais acessível essa interpretação.

²⁶ A razão do Estado do Amapá se encontrar nesse grupo se dá pela maior participação das variáveis associadas, principalmente o capital humano, quando comparado aos demais estados das Regiões Norte e Nordeste.

Figura 5

Índice de Desenvolvimento Educacional médio para os estados brasileiros e para o Brasil — 2007-11



6 Conclusões

A questão da educação nacional é um tema que ocupa a pauta de debates há décadas, no Brasil. Com o passar do tempo, novas questões vão surgindo, e poucas incógnitas inerentes ao seu desempenho são solucionadas. Assim, a utilização da estatística multivariada, particularmente, da análise fatorial, possibilitou a criação do Índice de Desenvolvimento Educacional. A sua importância está na capacidade de efetuar uma agregação dos 17 indicadores de desempenho da educação em cada um dos estados brasileiros, em uma única variável, para os cinco anos de análise. Desse modo, com o objetivo central sendo a criação desse índice e a análise de alguns dos indicadores considerados pontos cruciais, o desenvolvimento da educação é representado através do IDE.

Inicialmente, a análise de alguns indicadores já serve para elucidar o desenvolvimento heterogêneo apresentado pelo sistema educacional brasi-

leiro, já que, mesmo tendo passado por um processo de reorganização nas últimas décadas, as resultantes são inferiores às desejadas. Nesse âmbito, é frequente encontrar traços na literatura que apontam que o baixo desempenho educacional brasileiro acarreta à nação vários problemas sociais e econômicos. Pelo lado social, o baixo nível educacional do brasileiro é considerado um dos grandes causadores da pobreza e da má distribuição de renda no País. Pelo lado econômico, prejudica o crescimento da economia e aumenta o atraso tecnológico em relação às outras economias.

Desse modo, os resultados provenientes da presente pesquisa apontam que estados e regiões com desenvolvimento econômico em estágios mais elevados que os demais apresentam valores superiores para o IDE. Assim, o topo do *ranking* é ocupado pelo Distrito Federal, o qual apresenta melhor desempenho em vários dos indicadores de análise, enquanto a maior concentração dos estados se dá na categoria de médio desenvolvimento.

Do lado contrário, na cauda inferior do desempenho da educação, estão os estados da Região Nordeste, a qual também se encontra em um diferente estágio quando comparada, na esfera econômica, às outras regiões. Desse modo, ao identificar isso, as políticas governamentais podem apresentar maior grau de eficiência nessa região.

Face às estatísticas que foram apresentadas anteriormente, é de suma importância melhorar o sistema de ensino brasileiro. Um dos primeiros passos para isso é garantir um conhecimento detalhado sobre a educação brasileira e sua evolução recente em diferentes recortes temporais e geográficos. Assim, o que aqui é discutido contempla indicadores que possam possibilitar, de alguma forma, o entendimento e uma possível evolução dessa situação.

Por fim, recomendam-se futuras pesquisas, utilizando o IDE para avaliar a situação da educação, em nível municipal, de estados específicos, bem como, através dos aspectos teóricos da literatura recente desenvolvida por Hanushek e Woessmann, podem-se incorporar novas medidas de avaliação do desempenho da educação (por exemplo, ENEM, PISA e SAEB). Além disso, dentre as limitações do presente estudo, pode-se argumentar que o número de indicadores utilizados na extração das cargas fatoriais é pequeno, sendo que podem ser obtidos resultados mais robustos com inclusão de novas variáveis.

Referências

ACHILLES, C. **Let's Put Kids First, Finally: Getting Class Size Right.** Thousand Oaks, CA: Corwin, 1999.

AFONSO, A.; AUBYN, M. St. Cross-country efficiency of secondary education provision: A semi-parametric analysis with non-discretionary inputs. **Economic modeling**, [S.l.], v. 23, n. 3, p. 476-491, 2006.

AGHION, P. *et al.* The causal impact of education on economic growth: evidence from US. 2009. **Brookings Papers on Economic Activity**, Washington, DC, spring 2009. Disponível em: <http://www.brookings.edu/economics/bpea/~media/Files/Programs/ES/BP_EA/2009_spring_bpea_papers/2009_spring_bpea_aghion_et_al.pdf>. Acesso em: 28 abr. 2016.

ALTINOK, N. **Human capital quality and economic growth**. Dijon: Institute for Research in the Sociology and Economics of Education, 2007. (Working paper DT, 2007/1).

ANDERSON, R. *et al.* **Análise Multivariada de Dados**. 6. ed. Porto Alegre: Bookman, 2009.

BALDACCI, E. *et al.* Social Spending, Human Capital, and Growth in Developing Countries. **World Development**, [S.l.], v. 36, n. 8, p. 1317–1341, 2008.

BARBER, M.; MOURSHED, M. **How the world's best-performing school systems come out on top**. [S.l.]: McKinsey & Company, 2007.

BARRO, R. J. Economic growth in a cross section of countries. **The Quarterly Journal of Economics**, Cambridge, MA, v. 106, n. 2, p. 407-443, 1991.

BARRO, R. J. Human capital and growth. **American Economic Review**, Pittsburgh, PA, v. 91, n. 2, p. 12-17, 2001.

BARRO, R.; LEE, J. W. International data on educational attainment: updates and implications. **Oxford Economic Papers**, Oxford, v. 53, n. 3, p. 541-563, 2001.

BARROS, R. P. *et al.* Determinantes do desempenho educacional no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 31, n. 1, p. 1-42, 2001.

BLALOCK, H. M. **Measurement in the social sciences: Theories and strategies**. Chicago, IL: Aldine, 1974.

BOSWORTH, B. P.; COLLINS, S. M. The empirics of growth: An update. **Brookings Papers on Economic Activity**, Washington, DC, v. 34, n. 2, p. 113-206, 2003.

CAMPBELL, R.; HOMBO, C. M.; MAZZEO, J. **Trends in Academic Progress: Three Decades of Student Performance**. Washington, DC: U.S. Department of Education, 2000.

CASCIO, E.; CLARK, D.; GORDON, N. **Education and the age profile of literacy into adulthood**. Cambridge, MA: National Bureau of Economic Research, 2008. (NBER Working Paper, 14073)

CASTRO, J. A. Evolution and inequality in Brazilian education. **Educação & Sociedade**, Campinas, v. 30, n. 108, p. 673-697, 2009.

CHAY, K. Y.; McEWAN, P. J.; URQUIOLA, M. The central role of noise in evaluating interventions that use test scores to rank schools. **American Economic Review**, Pittsburgh, PA, v. 95, n. 4, p. 1237-1258, 2005.

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.

DOURADO, L. F. Políticas e Gestão da Educação Básica no Brasil: Limites e Perspectivas. **Educação e Sociedade**, Campinas, v. 28, n. 100, p. 921-946, out. 2007. Disponível em: <<http://www.cedes.unicamp.br>>. Acesso em: out. 2010.

FERNANDES, R.; GREMAUD, A. P. Qualidade da educação: avaliação, indicadores e metas. In: GIAMBIAGI, F. **Educação Básica no Brasil: construindo o país do futuro**. Rio de Janeiro: Elsevier, 2009. v. 1, p. 213-238.

FERREIRA JÚNIOR, S.; BAPTISTA, A. J. M. S.; LIMA, J. E. A modernização agropecuária nas microrregiões do Estado de Minas Gerais. **Revista de Economia e Sociologia Rural**, Brasília DF, v. 42, n. 1, p. 73-89, 2004.

FREITAS, C. A.; PAZ, M. V.; NICOLA, D. S. Analisando a modernização da agropecuária gaúcha: uma aplicação da análise fatorial e cluster. **Análise Econômica**, Porto Alegre, v. 25, n. 47, p. 121-149, 2007.

FUNDO NACIONAL DE DESENVOLVIMENTO DA EDUCAÇÃO (Brasil) (FNDE). **Sistema de Informações sobre Orçamentos Públicos em Educação (SIOPE): Relatório de Indicadores**. 2012. Disponível em: <<https://www.fnde.gov.br/siope/consultaIndicadoresEstaduais.do>>. Acesso em: dez. 2012.

GADOTTI, M. **Perspectivas atuais da educação**. Porto Alegre: Artes Médicas Sul, 2000.

GLEWWE, P. W. *et al.* School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010. Cambridge, MA: National Bureau of Economic Research, 2011. (NBER Working Paper, 17554).

GUNDLACH, E.; WOESSMANN, L.; GMELIN, J. The decline of schooling productivity in OECD countries. **The Economic Journal**, Malden, MA, v. 111, n. 471, p. 135-147, 2001.

HADDAD, S. **O direito à educação no Brasil**. São Paulo: DhESC Brasil, 2003. Disponível em: <<http://escoladegestores.mec.gov.br/site/8-biblioteca/pdf/sergiohaddad.pdf>>. Acesso em: 6 maio 2016.

HAIR, J. F. *et al.* **Multivariate data analysis**. Upper Saddle River, NJ: Pearson Prentice Hall, 2006. v. 6.

HALL, R.; JONES, C. **Why do some countries produce so much more output per worker than others?** Cambridge, MA: National Bureau of Economic Research, 1998. (NBER Working Paper, 6564).

HANUSHEK, E. A.; LUQUE, J. A. Efficiency and equity in schools around the world. **Economics of Education Review**, [S.l.], v. 22, n. 5, p. 481-502, 2003.

HANUSHEK, E. A.; WOESSMANN, L. Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. **Journal of Economic Growth**, [S.l.], v. 17, n. 4, p. 267-321, 2012.

HANUSHEK, E. A.; WOESSMANN, L. Education and economic growth. **Education Next**, Cambridge, MA, v. 8, n. 2, p. 62-70, 2008.

HANUSHEK, E. A.; WOESSMANN, L. **Schooling, cognitive skills, and the Latin American growth puzzle**. Cambridge, MA: National Bureau of Economic Research, 2009. (NBER Working Paper, 15066).

HANUSHEK, E. A.; WOESSMANN, L. **The role of education quality for economic growth**. [S.l.]: World Bank, 2007. (Policy Research Working Paper, n. 4122).

HANUSHEK, E. A.; WOESSMANN, L. **The economics of international differences in educational achievement**. Cambridge, MA: National Bureau of Economic Research, 2010. (NBER Working Paper, 15949).

HANUSHEK, E.; RAYMOND, M. Improving educational quality: How best to evaluate our schools? In: KODRZYCKI, Y. (Ed.). **Education in the 21st Century: Meeting the Challenges of a Changing World**. Boston, MA: Federal Reserve Bank of Boston, 2003. p. 193-224.

HOFFMANN, R. **Componentes Principais e Análise Fatorial**. Piracicaba: Escola Superior de Agricultura Luiz de Queiroz, 1999. (Série Didática, n. 90).

INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). **Ipeadata: Dados sociais**. 2012. Disponível em <<http://www.ipeadata.gov.br>>. Acesso em: dez. 2012.

INSTITUTO DE PESQUISA ECONOMICA E SOCIAL (IPEA). **Brasil: o estado de uma nação**. Brasília, DF: IPEA, 2006.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Informações Estatísticas**. 2005. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 12 dez. 2012.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Informações Estatísticas**. 2006. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 12 dez. 2012.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Informações Estatísticas**. 2007. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 12 dez. 2012.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Informações Estatísticas**. 2008. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 12 dez. 2012.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Informações Estatísticas**. 2009. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 12 dez. 2012.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Informações Estatísticas**. 2010. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 12 dez. 2012.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Informações Estatísticas**. 2011. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 12 dez. 2012.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Informações Estatísticas**. 2012. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 12 dez. 2012.

JAMISON, E. A.; JAMISON, D. T.; HANUSHEK, E. A. The effects of education quality on mortality decline and income growth. **Economics of Education Review**, [S.l.], v. 26, n. 6, p. 772-789, 2007.

KLEFENS, P. C. O. **O biplot da análise fatorial multivariada**. 2009. 103 f. Dissertação (Mestrado) - Universidade de São Paulo, Piracicaba, 2009.

LEE, D. W.; LEE, T. H. Human capital and economic growth: Tests based on the international evaluation of educational achievement. **Economics Letters**, [S.l.], v. 47, n. 2, p. 219-225, 1995.

LEVINE, R.; RENELT, D. A Sensitivity Analysis of Cross-Country Growth Regressions. **American Economic Review**, Pittsburgh, PA, v. 82, n. 4, p. 942-963, 1992.

MANKIW, N. G.; ROMER, D.; WEIL, D. A Contribution to the Empirics of Economic Growth. **The Quarterly Journal of Economics**, Cambridge, MA, v. 107, n. 2, p. 407-437, 1992.

MANLY, B. F. J. **Multivariate statistical methods-a primer**. New York: Chapman and Hall, 1986.

MENEZES FILHO, N. A. **A Evolução da Educação no Brasil e seu Impacto no Mercado de Trabalho**. São Paulo: Departamento de Economia/USP, 2001. Disponível em: <www.ifb.com.br/documentos/artigo_naercio.pdf>. Acesso em: 6 maio 2016.

MINGOTI, S. **Análise de dados através de métodos de estatística multivariada: Uma Abordagem Aplicada**. Belo Horizonte: Editora UFMG, 2005.

MONTEIRO, V. P.; PINHEIRO, J. C. Critério para implantação de tecnologias de suprimentos de água potável em municípios cearenses afetados pelo alto teor de sal. **Revista de Economia e Sociologia Rural**, Rio de Janeiro, v. 42, n. 2, p. 365-387, 2004.

OLIVEIRA, D. A. Profissão docente e gestão democrática da educação. **Revista Extra-classe**, [S.l.], v. 1. p. 210-217, 2009.

OLIVEIRA, R. P. Da universalização do ensino fundamental ao desafio da qualidade: uma análise histórica. **Educação e Sociedade**, Campinas, v. 28, n. 100, p. 661-690, 2007.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). **Education at a Glance 2014: OECD Indicators**. [S.l.], 2014.

PARO, V. H. Educação para a democracia: o elemento que falta na discussão da qualidade do ensino. **Revista Portuguesa de Educação**, Braga, v. 13, n. 1, p. 23-38, 2000.

PODRECCA, E.; CARMELI, G. **Education and Growth: A dynamic Analysis with Panel Data**. Fisciano: DiSES, 2004. (Working Papers DiSES).

RIANI, J. L. R.; GOLGHER, A. B. Indicadores educacionais confeccionados a partir de bases de dados IBGE. In: RIOS-NETO, E. L. G.; RIANI, J. de L. R. (Org.). **Introdução à Demografia da Educação**. Campinas: ABEP, 2004. v. 1, p. 89-128.

RIGOTTI, J. I. R. Variáveis de educação dos censos demográficos brasileiros de 1960 a 2000. In: RIOS-NETO, E. L. G.; RIANI, J. de L. R. (Org.). **Introdução à demografia da educação**. Campinas: ABEP, 2004. v. 1, p. 129-142.

RIGOTTI, J. I. R.; CERQUEIRA, C. A. As bases de dados do INEP e os indicadores educacionais: conceitos e aplicações. In: RIOS-NETO, E. L. G.; RIANI, J. de L. R. (Org.). **Introdução à demografia da educação**. Campinas: ABEP, 2004. v. 1, p. 73-88.

SAHLGREN, G. H. **Incentive to Invest?** How education affects economic growth. London: Adam Smith Institute, 2014. Disponível em: <<http://www.adamsmith.org/wpcontent/uploads/2014/07/IncentivetolInvest.pdf>>. Acesso em: 6 maio 2016.

SALA-I-MARTÍN, X.; DOPPELHOFER, G.; MILLER, R. I. Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. **American Economic Review**, Pittsburgh, PA, v. 94, n. 4, p. 813-835, 2004.

SANDER, B. **Administração da Educação no Brasil: genealogia do conhecimento**. Brasília, DF: Liber, 2007.

SCHANZENBACH, D. W. **Does Class Size Matter?** Boulder, CO: Education Policy Center, 2014.

SOUZA, N. J.; OLIVEIRA, J. C. Relações entre geração de conhecimento e desenvolvimento econômico. **Revista Análise**, Porto Alegre, v. 17, n. 2, p. 211-223, 2006.

SPAULL, N. **South Africa's education crisis**: The quality of education in South Africa 1994-2011. Johannesburg: Centre for Development and Enterprise, 2013.

SPEARMAN, C. "General intelligence" objectively determined and measured. **The American Journal of Psychology**, Champaign, IL, v. 15, n. 2, p. 201-293, 1904.

STEVENS, P.; WEALE, M. **Education and economic growth**. London: National Institute of Economic and Social Research, 2003. Disponível em: <http://cee.lse.ac.uk/conference_papers/28_11_2003/martin_weale.pdf>. Acesso: 6 maio 2016.

TABACHNICK, B.; FIDELL, L. **Using multivariate analysis**. Needham Heights: Allyn & Bacon, 2007.

THURSTONE, L. L. Multiple factor analysis. **Psychological Review**, Washington, DC, v. 38, n. 5, p. 406-427, 1931.

UNITED NATIONS EDUCATIONAL SCIENTIFIC AND CULTURAL ORGANIZATION (UNESCO). **Monitoramento dos Objetivos de Educação para Todos no Brasil**. 2010. Disponível em: <<http://unesdoc.unesco.org/images/0018/001899/189923por.pdf>>. Acesso em: 6 maio 2016.

UNITED NATIONS EDUCATIONAL SCIENTIFIC AND CULTURAL ORGANIZATION (UNESCO). **Relatório de Monitoramento Global de EPT**. Paris, 2013. Disponível em: <<http://unesdoc.unesco.org/images/0022/002256/225654por.pdf>>. Acesso em: 6 maio 2016.

WOESSMANN, L.; WEST, M. Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. **European Economic Review**, [S.l.], v. 50, n. 3, p. 695-736, 2006.

ZAMBRANO, C.; LIMA, J. E. Análise Estatística Multivariada de Dados Socioeconômicos. In: SANTOS, M. L.; VIEIRA, W. C. (Ed.). **Métodos quantitativos em economia**. Viçosa: UFV, 2004. p. 556-577.

ZELLER, R. A.; CARMINES, E. G. **Measurement in the social sciences**: The link between theory and data. Cambridge: Cambridge University Press, 1980.

