

Parte II

Bases de dados educacionais

Capítulo 1

As bases de dados do INEP e os indicadores educacionais: conceitos e aplicações

José Irineu Rangel Rigotti e Cézár Augusto Cerqueira

Capítulo 2

Indicadores educacionais confeccionados a partir de bases de dados do IBGE

Juliana de Lucena Ruas Riani e André Braz Golgher

Capítulo 3

Variáveis de educação dos censos demográficos brasileiros de 1960 a 2000

José Irineu Rangel Rigotti

Capítulo 2:

INDICADORES EDUCACIONAIS CONFECCIONADOS A PARTIR DE BASES DE DADOS DO IBGE

Juliana de Lucena Ruas Riani*
André Braz Golgher**

Este capítulo introduz alguns aspectos das bases domiciliares fornecidas pelo IBGE – os Censos Demográficos e as Pesquisas Nacionais de Amostra por Domicílios (PNADs) –, analisando as suas vantagens e desvantagens. Estas bases são abordadas principalmente no que tange às variáveis educacionais nelas presentes. Tanto os Censos quanto as PNADs são importantes para o estudo da educação por possibilitarem o cálculo de vários indicadores educacionais que podem ser desagregados em vários níveis de unidades geográficas e em variados grupos demográficos e/ou socioeconômicos, como nível de renda, idade, sexo, grupo educacional dos pais etc., além de permitirem o estudo de séries históricas. Num segundo momento, é apresentada a metodologia de cálculo para um conjunto de indicadores educacionais derivados a partir das bases domiciliares. Com estes indicadores foi possível construir um quadro da evolução da educação brasileira, bem como analisar vários aspectos e problemas do ensino no país.

No Brasil existem duas importantes bases de dados domiciliares, que são os Censos Demográficos e as Pesquisas Nacionais por Amostra de Domicílios (PNADs), ambas organizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Elas possuem estruturas semelhantes, porém o Censo Demográfico é bem mais amplo, abrangendo toda a população, ao passo que as PNADs são mais restritas, compreendendo apenas uma parcela da população (amostra).

Neste capítulo abordaremos, inicialmente, as principais características dessas duas bases e as vantagens e desvantagens na utilização de cada uma delas. Em um segundo momento serão analisados alguns indicadores gerados a partir dessas bases, tais como taxa de analfabetismo, anos médios de estudo,

*Doutoranda em Demografia e pesquisadora no Centro de Desenvolvimento e Planejamento Regional (Cedeplar) da Universidade Federal de Minas Gerais (UFMG).

**Doutor em Demografia pelo Cedeplar/UFMG e professor do Departamento de Economia da Faculdade de Administração e Ciências Econômicas (FACE) da UFMG.

taxa de atendimento, taxa de escolaridade líquida e bruta, taxa de distorção idade-série etc.

Censos Demográficos

Segundo o IBGE, a primeira contagem da população realizada no Brasil ocorreu em 1872. A partir de 1890 as contagens populacionais tornaram-se atividades regulares, com periodicidade decenal. Em 1940 elas passaram a seguir as recomendações dos organismos internacionais, sendo o levantamento desse ano considerado o primeiro Censo moderno do país.

Os Censos Demográficos brasileiros vêm se aprimorando cada vez mais e atualmente podem ser considerados bem completos, possibilitando o estudo da realidade nacional sob vários aspectos. Com essa base pode-se obter, dentre outras, informações sobre sexo, idade, condição domiciliar, cor, estado civil, migração, educação, rendimento, trabalho, fecundidade, nupcialidade etc. do indivíduo. Os dados permitem também analisar características domiciliares tais como abastecimento de água, destino do lixo, tipo de escoadouro, uso da instalação sanitária, existência de eletrodomésticos, número de cômodos etc. Essas informações têm subsidiado a confecção e implantação de políticas públicas com maior eficácia em diversas áreas, como educação, saúde, habitação, saneamento, migração, emprego e rendas.

O Censo Demográfico é uma pesquisa complexa e com altos custos. O de 2000, por exemplo, envolveu 200 mil pessoas no processo de coleta dos dados, que visitaram mais de 54 milhões de domicílios, e mais de 30 mil supervisores. Devido a essa complexidade, o Censo só é realizado a cada dez anos, o que às vezes o coloca em grande desvantagem com relação às PNADs, principalmente nos anos mais afastados da data da pesquisa.

Normalmente, os Censos Demográficos são realizados no segundo semestre do ano. O período de coleta de dados do Censo de 2000 foi do dia 1º de agosto – a data de referência – até o dia 30 de novembro. No Censo de 1991, a data de referência foi 1º de setembro.

As informações contidas nessa base de dados têm origem em dois questionários. Um questionário básico, aplicado a todos os domicílios, contém informações sobre as características básicas da população, das pessoas responsáveis pelos domicílios, e dos próprios domicílios. O outro questionário é bem mais amplo, englobando, além das questões contidas no primeiro, outras referentes a religião, cor ou raça, deficiência, migração, escolaridade, fecundidade, nupcialidade, trabalho, rendimento e outras características domiciliares. Este

questionário é aplicado apenas a uma parcela dos domicílios, resultando em uma amostra que pode ser expandida por meio de pesos calculados para cada domicílio pesquisado e seus respectivos moradores fornecidos pelo IBGE¹.

Esses dois questionários censitários permitem estudar vários aspectos da realidade brasileira com uma riqueza muito grande de informação. O IBGE disponibiliza várias tabelas geradas a partir dos dados censitários por meio de publicações e via *internet*, pelo endereço eletrônico www.ibge.gov.br, ou através do Sistema IBGE de Recuperação Automática (SIDRA), no *site* www.sidra.ibge.gov.br.

As informações específicas sobre educação do questionário básico (resultado do universo), à disposição nos *sites* do IBGE, restringem-se a dados sobre alfabetização. Contudo, estão disponíveis para *download* muitos outros dados obtidos a partir do questionário mais detalhado, como, por exemplo, informações sobre anos de estudo, freqüência por nível escolar e freqüência na rede pública ou privada, classificadas por sexo e por grupos de idade. Esses dados podem ser analisados para Brasil, Grandes Regiões ou Unidades da Federação. A partir deles são calculados alguns indicadores, como a taxa de analfabetismo, a taxa de escolarização bruta e líquida, a taxa de freqüência à escola, a proporção da população com *x* anos de estudo etc. A estimação desses indicadores será um dos itens abordados mais adiante.

Todavia, quando se deseja obter um maior refinamento destas taxas, como, por exemplo, calcular a taxa de analfabetismo por cor, por faixa de renda domiciliar, por escolaridade dos pais etc., ou quando se pretende calcular outros indicadores mais específicos, pode-se solicitar ao IBGE tabulações especiais ou adquirir o CD com os microdados da amostra, que permitem gerar tabulações especiais próprias.

Esses últimos dados constituem o menor nível de desagregação. Nesses arquivos, a unidade geográfica mínima de análise são os municípios, podendo chegar ao nível de distrito, subdistrito ou, em alguns casos, a áreas ainda mais desagregadas. Eles são importantes por possibilitarem vários cruzamentos entre as características individuais, domiciliares e socioeconômicas dos estudantes e, desta forma, permitirem a análise dos principais determinantes do desempenho educacional, fornecendo importantes subsídios para a formulação de políticas públicas educacionais. Posteriormente, serão descritos alguns aspectos metodológicos relativos à obtenção dessas tabulações especiais.

¹ Para mais informações, veja a documentação dos microdados da amostra do Censo de 2000 em IBGE (2002).

Pesquisas Nacionais por Amostra de Domicílios

As Pesquisas Nacionais por Amostra de Domicílios (PNADs), como o próprio nome diz, são pesquisas realizadas em uma amostra de domicílios. Essa pesquisa apresenta algumas vantagens e desvantagens com relação ao Censo Demográfico. A escolha de qual base usar, ou mesmo o uso de ambas em um mesmo estudo, depende dos objetivos a serem alcançados.

A primeira PNAD foi realizada em 1967, com uma abrangência bem limitada. Aos poucos, seu universo foi sendo expandido, e desde 1981 a PNAD já cobre todo o Brasil, com exceção da área rural da região Norte². Esta seria a sua primeira desvantagem com relação ao Censo: a impossibilidade de realizar estudos sobre essa região específica, ou mesmo de incluí-la em estudos sobre o Brasil ou a região Norte. Por outro lado, a pesquisa tem periodicidade anual, não sendo realizada apenas em anos censitários³. Na verdade, ela surgiu da necessidade de um acompanhamento anual do quadro socioeconômico brasileiro, de contarmos com análises baseadas em dados mais recentes. Essa seria a principal vantagem das PNADs em relação aos Censos Demográficos. Entretanto, para que uma pesquisa anual fosse viável economicamente, a complexidade de coleta de dados deveria ser muito menor que a dos Censos. Para que isso fosse alcançado, mantendo-se uma estrutura de questionário semelhante à do Censo, optou-se por restringir a pesquisa a uma amostra bem menor que a do Censo, o que limita a representatividade das PNADs, cujos dados podem ser desagregados apenas para unidades da Federação e regiões metropolitanas. Esta seria a principal desvantagem das PNADs em relação aos Censos.

Com relação aos temas investigados, a pesquisa amostral é constituída de um levantamento básico, com questões referentes a aspectos gerais da população como educação, trabalho, rendimento e habitação, e de levantamentos suplementares que abordam diferentes temas em cada ano.

De 1981 a 1990, o questionário básico da PNAD manteve a mesma estrutura. Uma das poucas alterações no período foi a introdução dos quesitos sobre cor e sobre a existência de rádio e televisão nos domicílios, respectivamente em 1987 e 1988. No período de 1992 a 1999 foram feitas algumas mudanças mais significativas na estrutura do questionário, como a inclusão dos temas da migração, fecundidade e nupcialidade no levantamento básico, registrando-se também uma mudança metodológica nos quesitos referentes a educação e

² Tocantins é o único estado dessa região cuja área rural é coberta pela pesquisa.

³ Em 1994, por motivos técnicos, não foi realizado o levantamento da PNAD.

trabalho. Dessa forma, quando se deseja fazer uma análise da série histórica utilizando diversas PNADs, deve-se ter o cuidado de considerar essas mudanças metodológicas, bem como as revisões feitas nas projeções populacionais usadas na ampliação da amostra.

Os temas investigados nos levantamentos suplementares permitem realizar pesquisas mais detalhadas sobre alguns pontos específicos que não seriam possíveis apenas a partir do questionário básico das PNADs. Os temas investigados nos levantamentos suplementares das PNADs, desde a década de 1980, foram: saúde (1981 e 1998); participação político-social e estoque de aparelhos utilizadores de energia (1982); educação (1982); mão-de-obra e previdência (1983); fecundidade feminina (1984); situação do menor (1985); anticoncepção, acesso a serviço social, suplementação alimentar e associativismo (1986); mercado de trabalho (1989 e 1999); mobilidade social (1996); trabalho infantil e de adolescentes (2001).

Os dados das PNADs também são disponibilizados pelo IBGE via publicações e *internet*. Especificamente sobre educação, encontram-se dados sobre o número de pessoas classificadas de diferentes maneiras: analfabetos por sexo e situação de domicílio; por anos de estudo, por sexo e por situação de domicílio; por grau e série que frequentavam, por sexo, por situação de domicílio e por grupos de idades. Estes dados estão disponíveis para Brasil, Grandes Regiões e Unidades da Federação. Assim como os dados censitários, os dados das PNADs obtidos em publicações ou via *internet* permitem calcular vários indicadores educacionais. Para uma análise mais refinada, porém, deve-se adquirir os microdados.

A utilização de microdados em demografia da educação

Os microdados dos Censos e das PNADs são semelhantes em formato e apresentam o menor nível de desagregação possível. Eles fornecem os resultados em dois arquivos distintos. Um contendo as informações para pessoas e outro, para domicílios. Como pode ser visto pela Tabela 1, que apresenta um exemplo fictício, cada linha do banco de dados é um indivíduo diferente e cada coluna ou grupo de colunas representa um dos quesitos investigados.

No exemplo da Tabela 1, que não corresponde exatamente ao observado na realidade, mas que serve para facilitar o entendimento do leitor, as primeiras colunas indicam qual é o indivíduo de que estamos tratando, o segundo grupo de colunas mostra a idade de cada indivíduo, e o terceiro grupo apresenta o sexo de forma numérica, 1 para homens e 3 para mulheres. Nos dados reais existem muitas outras variáveis, mas a matriz abaixo pode ser considerada um fragmento

desses dados. Observam-se três pessoas distintas. O indivíduo de número 0111 é um homem com idade de 24 anos; o segundo, uma mulher de 20 anos; o terceiro, uma criança de 3 anos do sexo masculino.

TABELA 1
Exemplo fictício de microdados para o arquivo Pessoas

Variável							
Indivíduo				Idade			Sexo
0	1	1	1	0	2	4	1
0	1	1	2	0	2	0	3
0	1	1	3	0	0	3	1

O arquivo de domicílios é semelhante ao de pessoas, ou seja, cada linha da base de dados refere-se a um domicílio diferente. Estes dois arquivos podem, inclusive, ser unidos, permitindo análises ainda mais ricas.

Os quesitos referentes à educação existentes nos microdados possibilitam calcular os anos de estudo da população, bem como o grau e a série dos que estão freqüentando a escola em diferentes grupos da população brasileira. Por exemplo, se um pesquisador está interessado em saber qual a proporção de mulheres com filhos com idade entre 15 e 19 anos que não estão freqüentando a escola nem estão trabalhando no município de Jaboatão, basta ele cruzar todas essas variáveis em uma análise direta dos microdados. A seguir veremos mais detalhadamente as variáveis educacionais presentes nos quesitos das pesquisas domiciliares, inicialmente as dos Censos e, em seguida, as das PNADs.

Cabe ressaltar que nos microdados do IBGE existem variáveis derivadas pelo próprio Instituto. Uma delas é a classificação da população por anos de estudo. Porém, existem algumas imperfeições nessa variável, como o agrupamento dos anos de estudo mais avançados em uma única categoria, ou a contagem de mais um ano de estudo para alunos de cursos com duração maior que a definida como padrão. Isso ocorre, por exemplo, no Ensino Médio, onde a duração padrão dos cursos é de três anos, mas há cursos com quatro anos de duração. No cálculo dos anos de estudo do IBGE, o indivíduo que freqüentou este último tipo de curso terá 12 anos de estudo, porém muitos analistas consideram que esta pessoa tem apenas 11 anos de estudo. Por esse motivo, alguns pesquisadores preferem criar sua própria variável "anos de estudo", a partir dos demais quesitos existentes na base de dados.

Uma vez discutidas as características gerais dos Censos e PNADs, passemos aos quesitos dos microdados referentes à educação presentes nessas bases de dados. O objetivo desta apresentação é, além de familiarizar o leitor com o tipo

de informação existente nas pesquisas domiciliares brasileiras, demonstrar a grande riqueza de interpretações possíveis pelo cruzamento das diversas variáveis.

Ambas as pesquisas, como já vimos, vêm sofrendo mudanças constantes. Essas reformulações são tanto de ordem metodológica como visam à inclusão/exclusão de quesitos. Apesar disso, essas bases de dados apresentam muitas informações iguais ou semelhantes que formam um núcleo comum.

Nas próximas seções apresentamos os quesitos educacionais presentes no Censo Demográfico mais recente e na última PNAD, respectivamente.

Quesitos do Censo de 2000 referentes à educação⁴

No Censo de 2000, os quesitos sobre educação são semelhantes aos encontrados nas PNADs da década de 90, mas em menor número. As variáveis educacionais nele incluídas são:

Variável v0428 => *Sabe ler e escrever*

Respostas possíveis: 01 - Sim
02 - Não

Variável v0429 => *Freqüenta escola ou creche*

Respostas possíveis: 01 - Sim, rede particular
02 - Sim, rede pública
03 - Não, já freqüentou
04 - Nunca freqüentou

Variável v0430 => *Curso que freqüenta*

Respostas possíveis: 01 - Creche
02 - Pré-Escolar
03 - Classe de Alfabetização
04 - Alfabetização de Adultos
05 - Ensino Fundamental ou 1º Grau – regular seriado
06 - Ensino Fundamental ou 1º Grau – regular não-seriado
07 - Supletivo (Ensino Fundamental ou 1º Grau)
08 - Ensino Médio ou 2º Grau – regular seriado

⁴ Na Parte 3, capítulo 2, deste livro é feita uma discussão mais detalhada das variáveis educacionais presentes nos Censos Demográficos desde 1960.

- 09 - Ensino Médio ou 2º Grau – regular não-seriado
- 10 - Supletivo (Ensino Médio ou 2º Grau)
- 11 - Pré-Vestibular
- 12 - Superior – Graduação
- 13 - Superior – Mestrado ou Doutorado

Variável v0431 => *Série que frequenta*

Respostas possíveis: 01 a 08 - Série que frequenta (1ª a 8ª série)

- 09 - Curso não-seriado

Variável v0432 => *Curso mais elevado que frequentou, tendo concluído ao menos uma série*

Respostas possíveis: 01 - Alfabetização de Adultos

- 02 - Antigo Primário
- 03 - Antigo Ginásio
- 04 - Antigo Clássico, Científico etc.
- 05 - Ensino Fundamental ou 1º Grau
- 06 - Ensino Médio ou 2º Grau
- 07 - Superior – Graduação
- 08 - Mestrado ou Doutorado
- 09 - Nenhum

Variável v0327 => *Última série concluída com aprovação*

Respostas possíveis: 01 a 08 - Série que concluiu (1ª a 8ª série)

- 09 - Curso não-seriado
- 10 - Nenhum

Variável v0434 => *Concluiu o curso no qual estudou*

Respostas possíveis: 01 - Sim

- 02 - Não

Variável v4355 => *Código do curso mais elevado concluído*

Respostas possíveis: as mesmas da v0432

Quesitos das PNADs referentes à educação

As PNADs sofreram mudanças significativas na metodologia de coleta dos dados referentes à educação. Pode-se considerar dois períodos em que os quesitos e a metodologia de coleta praticamente não se alteraram: de 1977 a 1990 e de 1992 a 2001. No primeiro período (1977-1990), havia cinco perguntas relacionadas à educação no levantamento básico, que só eram respondidas para os maiores de 5 anos de idade. A partir de 1995 esses quesitos foram expandidos para todas as idades. São eles:

Variável v311 => *Sabe ler e escrever*

Respostas possíveis: 01 - Sim

02 - Não

09 - Sem declaração

Variável v312 => *Série que freqüenta na escola*

Respostas possíveis: 00 - Outros/Não freqüenta

01 a 08 - Série que freqüenta (1ª a 8ª série)

09 - Sem declaração

Variável v314 => *Grau que freqüenta na escola*

Respostas possíveis: 00 - Não freqüenta

01 - Elementar

02 - Médio - 1º Ciclo

03 - Médio - 2º Ciclo

04 - 1º Grau

05 - 2º Grau

06 - Superior

07 - Pré-Escolar

08 - Alfabetização de Adultos

09 - Supletivo 1º Grau

10 - Supletivo 2º Grau

11 - Supletivo 1º Grau/Rádio e TV

12 - Supletivo 2º Grau/Rádio e TV

13 - Supletivo sem especificação

14 - Vestibular

15 - Doutorado/Mestrado

99 - Sem declaração - menor de 5 anos

Variável v315 => *Última série freqüentada com aprovação para pessoas que não mais freqüentam a escola*

Respostas possíveis: 00 - Ainda freqüenta escola

01 a 08 - Série que freqüentou (1ª a 8ª série)

09 - Sem declaração

Variável v317 => *Grau da última série freqüentada para pessoas que não mais freqüentam a escola*

Respostas possíveis: 00 - Ainda freqüenta escola

01 - Elementar

02 - Médio - 1º Ciclo

03 - Médio - 2º Ciclo

04 - 1º Grau

05 - 2º Grau

06 - Superior

07 - Doutorado/Mestrado

08 - Alfabetização de Adultos

99 - Sem declaração

As formas de utilização dessas variáveis também podem variar. Algumas vezes utiliza-se um único quesito de cada vez. Com a variável v311, por exemplo, é possível identificar o número de analfabetos, pois só é considerado analfabeto quem não sabe ler nem escrever um bilhete simples, e a partir desse dado pode-se calcular taxas de analfabetismo, como será descrito posteriormente. Na maioria dos estudos, entretanto, utiliza-se duas ou mais variáveis em conjunto. Por exemplo, com o cruzamento do grau e série das pessoas que freqüentam a escola é possível saber como a população de uma localidade se distribui entre os vários níveis escolares. Se incluirmos, ainda, informações sobre as pessoas que não freqüentam quaisquer estabelecimentos de ensino, pode-se classificar toda a população com relação aos anos de estudo do indivíduo.

Os quesitos listados acima são respondidos por todos, independentemente de o indivíduo freqüentar ou não a escola. Se a pessoa estivesse freqüentando a escola, teria de dar como resposta a opção "zero" (freqüenta escola) nas variáveis v315 e v317. Por outro lado, se não estivesse freqüentando escola, teria de responder "zero" nas variáveis v312 e v314. Esta forma de indagação não estruturada gerou algumas inconsistências, como a resposta simultânea e diferente de zero para as variáveis v314 e v317 ou para v312 e v315, ou a declaração de cursar alguma série e nenhum grau.

Para tentar diminuir alguns erros e inconsistências, a partir de 1992 foram introduzidas algumas mudanças na estrutura e metodologia das variáveis educacionais da PNAD, com a incorporação ao questionário de outras perguntas, agora indagadas de maneira estruturada. Esta parte sobre educação também foi expandida para todas as idades a partir de 1995, não mais se restringindo aos maiores de 5 anos, como nas pesquisas anteriores.

Os quesitos referentes à educação incluídos em 1992 e que constaram da última PNAD (2001)⁵ são:

Variável v0601 => *Sabe ler e escrever*

Respostas possíveis: 01 - Sim

03 - Não

09 - Ignorada

Variável v0602 => *Freqüenta escola ou creche*

Respostas possíveis: 02 - Sim

04 - Não

09 - Ignorada

Variável v0603 => *Qual o curso que freqüenta*

Respostas possíveis: 00 - Ignorado

01 - Regular de 1º Grau

02 - Regular de 2º Grau

03 - Supletivo de 1º Grau

04 - Supletivo de 2º Grau

05 - Superior

06 - Alfabetização de Adultos

07 - Creche

08 - Pré-Escolar

09 - Pré-Vestibular

10 - Mestrado ou Doutorado

Variável v0604 => *O curso que freqüenta é seriado*

Respostas possíveis: 02 - Sim

04 - Não

09 - Ignorada

⁵ Na PNAD de 2001 as variáveis v0603 e v0607 sofreram pequenas modificações nas suas categorias de resposta (ver dicionário dos microdados da amostra da PNAD de 2001).

Variável v0605 => *Qual a série que frequenta*

Respostas possíveis: 01 a 08 - Série que frequenta (1ª a 8ª série)
09 - Ignorada

Variável v0606 => *Anteriormente frequentou escola ou creche*

Respostas possíveis: 02 - Sim
04 - Não
09 - Ignorada

Variável v0607 => *Curso mais elevado que frequentou*

Respostas possíveis: 00 - Ignorado
01 - Elementar (Primário)
02 - Médio - 1º Ciclo (Ginásial)
03 - Médio - 2º Ciclo (Científico)
04 - 1º Grau
05 - 2º Grau
06 - Superior
07 - Mestrado ou Doutorado
08 - Alfabetização de Adultos
09 - Creche
10 - Pré-Escolar

Variável v0608 => *Este curso que frequentou era seriado*

Respostas possíveis: 02 - Sim
04 - Não
09 - Ignorada

Variável v0609 => *Concluiu com aprovação pelo menos a 1ª série deste curso que frequentou*

Respostas possíveis: 01 - Sim
03 - Não
09 - Ignorada

Variável v0610 => *Última série concluída neste curso que frequentou*

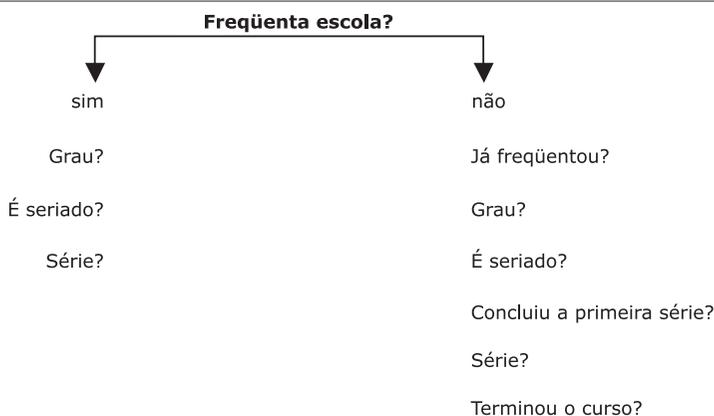
Respostas possíveis: 01 a 08 - Série que frequenta (1ª a 8ª série)
09 - Ignorada

Variável v0611 => *Concluiu este curso que frequentou anteriormente*

Respostas possíveis: 01 - Sim
03 - Não
09 - Ignorada

Para evitar os problemas de consistência observados nas PNADs realizadas até 1990, os quesitos passaram a ser respondidos de forma estruturada, como pode ser visto na Tabela 2. Ou seja, primeiro o indivíduo responde se frequenta a escola ou não. Caso a resposta seja positiva, o recenseado responde apenas às variáveis v0603, v0604 e v0605. Se a resposta for negativa, pula-se estas variáveis e passa-se para a v0606 ("Anteriormente frequentou escola ou creche?"), e se a resposta for positiva, passa-se para o restante dos quesitos.

TABELA 2
Estrutura de resposta dos quesitos de educação nas PNADs (1992 a 2001)



Fonte: Soares e Lima (2002).

As variáveis educacionais dos Censos e PNADs podem ser analisadas uma a uma, em grupos, ou em conjunto com os demais quesitos não diretamente relacionados com a educação, como sexo, idade, renda, se o domicílio possui eletricidade ou não etc. Os microdados e suas diversas variáveis, hoje disponíveis em CDs, são manipulados atualmente por microcomputadores comuns, por meio de programas computacionais ou algoritmos específicos.

Esses algoritmos são, de certa forma, bastante simples. A seguir é mostrado aquele que calcula a matrícula das pessoas que estão cursando a 1ª série do Ensino Fundamental tanto regular (seriado ou não) quanto supletivo.

se (v0430=5 e v0431=1) série=1.

se (v0430=6 e v0431=1) série=1.

se (v0430=7 e v0431=1) série=1.

Se, para estar na 1ª série, a pessoa deve estar cursando a série 1 (v0431=1) do curso 5, 6 ou 7, para saber o número de pessoas que estão cursando a 1ª

série basta cruzar estas duas variáveis, conforme o exemplo acima. Com a mesma lógica constrói-se a matrícula nas séries seguintes, obtendo-se a distribuição da população em todas as séries. A linguagem utilizada foi a do programa computacional SPSS.

No CD anexado a este livro encontram-se vários algoritmos para construir os anos de estudo tanto para as PNADs das décadas de 80 e 90, como para os Censos Demográficos de 1991 e 2000. Eles foram feitos utilizando a linguagem computacional SPSS, que é um pacote estatístico muito popular em ciências sociais, mas existem muitos outros pacotes que também podem ser utilizados.

Pela manipulação direta das bases de dados por meio de algoritmos específicos pode-se obter uma ampla gama de indicadores em demografia da educação. Na próxima seção são discutidos alguns deles.

Alguns indicadores de educação originados a partir das PNADs e Censos Demográficos

Com base nos Censos Demográficos e nas PNADs é possível calcular um conjunto de indicadores educacionais que possibilitam identificar os principais aspectos e problemas do sistema de ensino brasileiro. Pode-se, por exemplo, analisar as diferenças regionais com relação ao acesso à escola, atendimento e cobertura escolar, produtividade do ensino e disponibilidade de recursos educacionais. Lembramos, mais uma vez, que diferentes recortes geográficos podem ser feitos no cálculo dessas taxas, como para o Brasil, Grandes Regiões, Unidades da Federação e mesmo para municípios, esses últimos se a base de dados for algum Censo Demográfico.

Cabe ressaltar que o conjunto de indicadores educacionais é derivado da disponibilidade dos dados estatísticos, e que por meio deles pretende-se mensurar a realidade educacional do país tanto para fins de pesquisa acadêmica como de formulação de políticas públicas. A seguir são apresentadas as metodologias de cálculo e alguns dos resultados obtidos para alguns indicadores educacionais.

Taxa de analfabetismo

A taxa de analfabetismo é bastante utilizada por órgãos internacionais como um indicador que mede os níveis de desenvolvimento socioeconômico de países. É empregada, por exemplo, pela ONU na confecção dos Índices de Desenvolvimento Humano (IDH). Representa o quociente entre a população analfabeta e a população total de um mesmo grupo etário. Geralmente, calcula-se essa taxa para as pessoas com 10 anos ou mais de idade (mais utilizada pelo IBGE)

ou com 15 anos ou mais de idade (mais utilizada pelo INEP), mas existem inúmeras outras possibilidades. Considera-se analfabeto aquele indivíduo que é incapaz de ler e escrever ao menos um bilhete simples na sua língua de origem. Nos questionários das PNADs e dos Censos Demográficos ele é captado pela pergunta "Sabe ler e escrever?". Aquele que responde "não" é considerado analfabeto.

A fórmula abaixo mostra como calcular a taxa de analfabetismo:

$$TA = \left(\frac{P_{ana}}{P} \right) \times 100,$$

onde:

TA é taxa de analfabetismo;

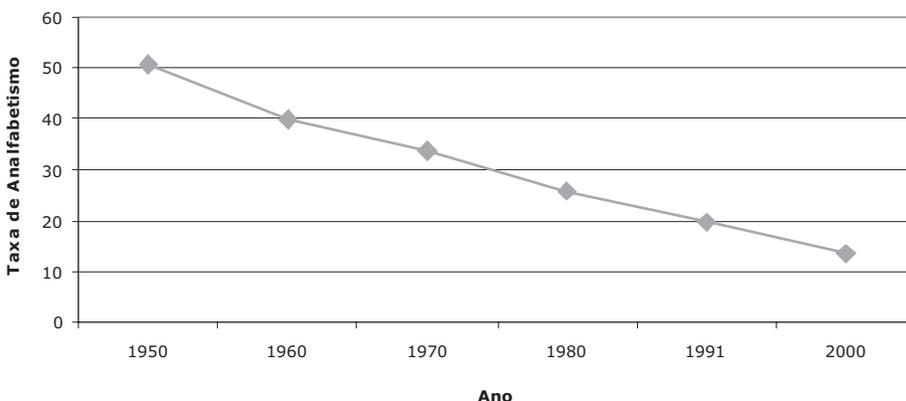
P_{ana} é a população analfabeta de um determinado grupo etário; e

P é a população total nesse mesmo grupo etário.

Como se sabe, a taxa de analfabetismo no Brasil apresentou uma significativa redução nas últimas décadas. No Gráfico 1 apresentamos a evolução da taxa de analfabetismo para a população brasileira de 15 anos ou mais no período de 1950 e 2000, utilizando dados dos Censos Demográficos. Percebe-se que esta taxa diminuiu muito na segunda metade do século XX: de 50% da população, nos anos 50, para aproximadamente 14% ao final do período.

GRÁFICO 1

Taxa de analfabetismo na faixa de 15 anos ou mais – Brasil, 1950/2000

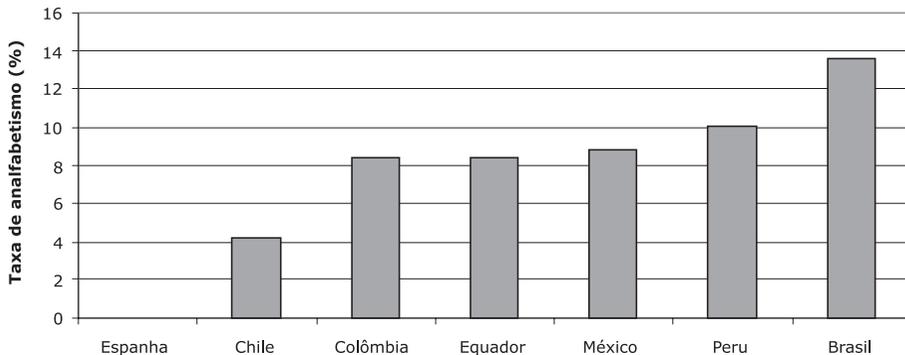


Fonte: IBGE, Censos Demográficos.

Apesar dessa marcante queda, a taxa de analfabetismo brasileira ainda apresenta um nível muito elevado quando comparada com a de outros países, inclusive países com nível de renda semelhante ao nosso. O Gráfico 2 compara os dados brasileiros com os de países selecionados, todos latinos e alguns com renda *per capita* inferior à nossa. Como se pode notar, a Espanha apresenta uma taxa de analfabetismo desprezível.

GRÁFICO 2

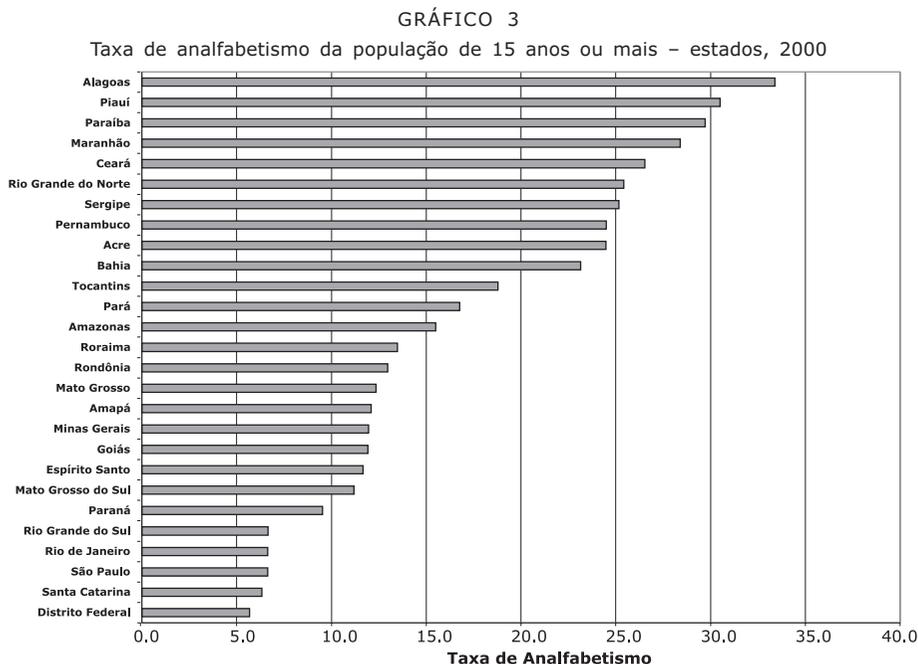
Taxa de analfabetismo da população de 15 anos ou mais para diferentes países – 2000



Fontes: PNUD e Unesco.

Além de exibir elevadas taxas de analfabetismo, o Brasil é um país extremamente heterogêneo. O Gráfico 3 mostra a grande desigualdade regional nessas taxas, apresentando os dados desagregados por estado. As taxas variam de 5,7%, no Distrito Federal, a 33,4% em Alagoas. De maneira geral, os estados das regiões Nordeste e Norte são os que apresentam as maiores taxas, o contrário ocorrendo nas regiões Sul e Sudeste. Grosso modo, os estados brasileiros podem ser classificados em três grupos distintos quanto ao nível dessa taxa: os de melhores índices, do Distrito Federal até o Rio Grande do Sul; os intermediários, do Paraná até Tocantins; e os que apresentam as maiores taxas, da Bahia em diante. Essa situação evidencia a necessidade de implementação de políticas públicas específicas voltadas para as áreas mais atrasadas em termos educacionais.

Quando, entretanto, a taxa de analfabetismo é calculada para toda a população acima de uma certa idade, como é o caso dos gráficos acima, referentes a pessoas com 15 anos ou mais de idade, ela engloba toda a história passada do



Fonte: IBGE, Censo Demográfico de 2000.

sistema de educação, ou seja, é uma medida de estoque. Desta forma, ela não seria apropriada para avaliar os recentes avanços na educação, uma vez que políticas educacionais geralmente são voltadas para a população em idade escolar, com prioridade para o sistema de ensino básico. Avanços educacionais recentes são, portanto, diluídos pelas deficiências do passado.

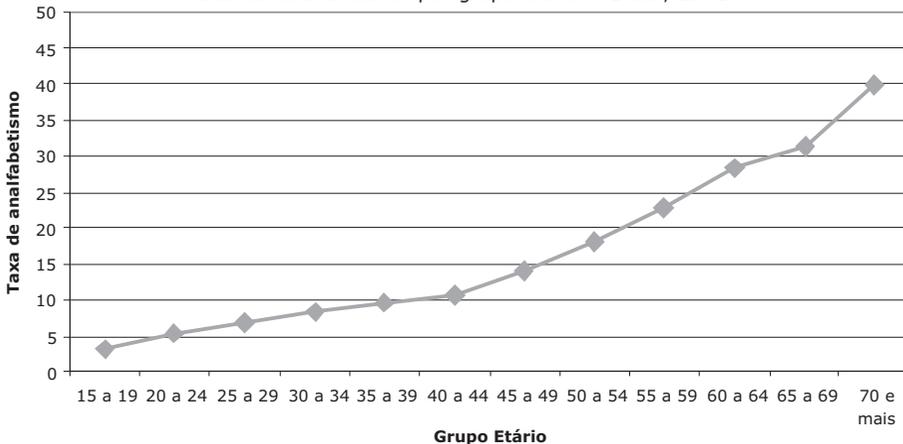
Para captar o avanço de políticas voltadas para o ensino básico e auxiliar na formulação de políticas de alfabetização de adultos, é conveniente analisar a taxa de analfabetismo por faixa etária. Desta forma, também se estaria tirando o efeito composição da estrutura etária que está embutido nesta taxa. Como o maior número de analfabetos concentra-se nas idades mais velhas, uma população mais envelhecida pode ter uma taxa de analfabetismo mais alta que outra mais jovem, mesmo que elas tenham as mesmas taxas por grupos etários.

No Gráfico 4 encontram-se as taxas de analfabetismo para diferentes grupos de idade em 2001. Note-se que, diferentemente dos Gráficos 1 e 3, que foram baseados nos dados do Censo Demográfico, o Gráfico 4 foi construído a partir dos resultados da PNAD. Fica bastante evidenciado o diferencial entre as gerações. A maior taxa nas populações mais idosas é resultado da deficiência do sistema

de ensino em anos passados. Os níveis próximos de zero para os grupos etários mais jovens captam o avanço na educação que vem ocorrendo de maneira gradativa e constante nas últimas décadas. Este quadro evidencia a importância de políticas de alfabetização de adultos para garantir um nível mínimo de dignidade e cidadania para a população mais idosa.

Além da possibilidade de desagregação das taxas por idade, pode-se refinar ainda mais o cálculo dessa taxa fragmentando a população analisada em grupos específicos, a partir de diferenças no sexo, na cor, no estrato de rendimento etc.

GRÁFICO 4
Taxa de analfabetismo por grupo etário – Brasil, 2001

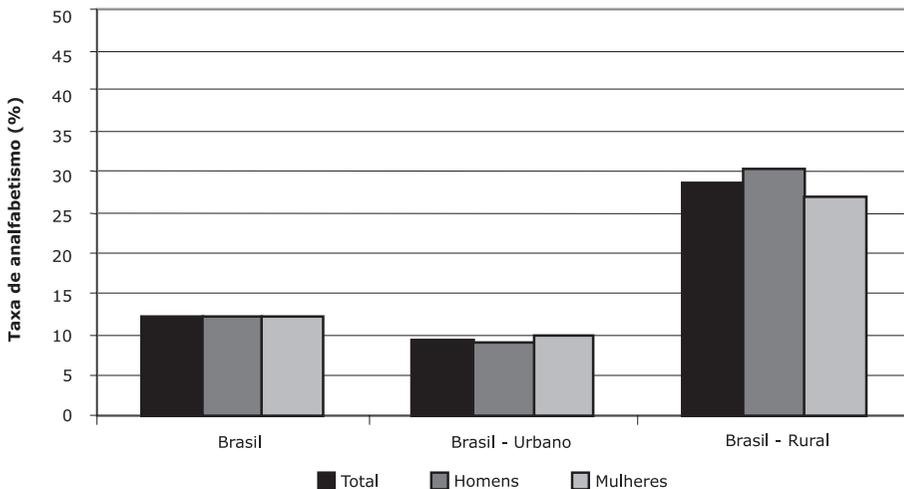


Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios 2001.

O Gráfico 5 é um exemplo de desagregação da taxa de analfabetismo das pessoas de 15 anos e mais por sexo e situação de domicílio. Percebe-se que há uma grande diferença entre os setores rural e urbano, sendo este último bem mais avançado em termos educacionais que o primeiro. Exemplificando, essa taxa em 2001 era por volta de 12% nas cidades, contra 30% no setor rural. Com relação ao sexo, verificam-se pequenas diferenças: no setor rural, as mulheres apresentavam uma maior taxa de analfabetismo, o contrário sendo observado no setor urbano.

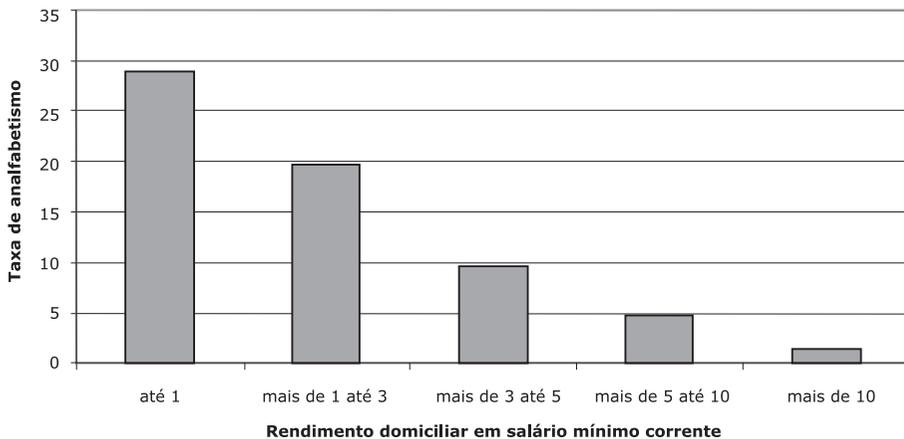
Outro exemplo de análise desagregada para a taxa de analfabetismo é apresentado no Gráfico 6, que traz a taxa de analfabetismo para diferentes estratos de renda domiciliar. O gráfico evidencia o que intuitivamente todos já sabem: que as classes mais pobres da população têm as maiores taxas de analfabetismo, sendo bem marcantes as diferenças entre os domicílios cujos chefes ganham até 1 salário mínimo e aqueles cujos chefes ganham mais de dez salários mínimos.

GRÁFICO 5
 Taxa de analfabetismo para a população de 15 anos ou mais por sexo e situação de domicílio – Brasil, 2001



Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios 2001.

GRÁFICO 6
 Taxa de analfabetismo para a população de 15 anos ou mais por rendimento domiciliar – Brasil, 2001



Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios 2001.

Anos médios de estudo ou escolaridade média

A escolaridade média corresponde à média dos anos de estudos concluídos por uma determinada população. Vale ressaltar que não se consideram os anos que a pessoa passou na escola, mas apenas aqueles em que ela teve aprovação. Como vimos, nas PNADs e Censos Demográficos constrói-se os anos de estudo com base nos quesitos sobre a série e o grau de quem freqüenta a escola ou já freqüentou um dia.

A fórmula abaixo exemplifica como são feitos os cálculos.

$$\text{Anos de estudos} = \frac{(0 \times P_0 + 1 \times P_1 + \dots + 17 \times P_{17})}{P}$$

onde:

P_0 é a população com zero ano de estudo;

P_1 é a população com um ano de estudo;

P_{17} é a população com 17 anos de estudo; e

P é a população total.

As pessoas com zero ano de estudo são aqueles que nunca estudaram e os que, se freqüentaram a escola, não conseguiram ser aprovados nem na 1ª série do Ensino Fundamental. Os indivíduos com um ano de estudo são aqueles que já foram aprovados na 1ª série do Ensino Fundamental mas ainda não foram aprovados na série seguinte, e assim sucessivamente.

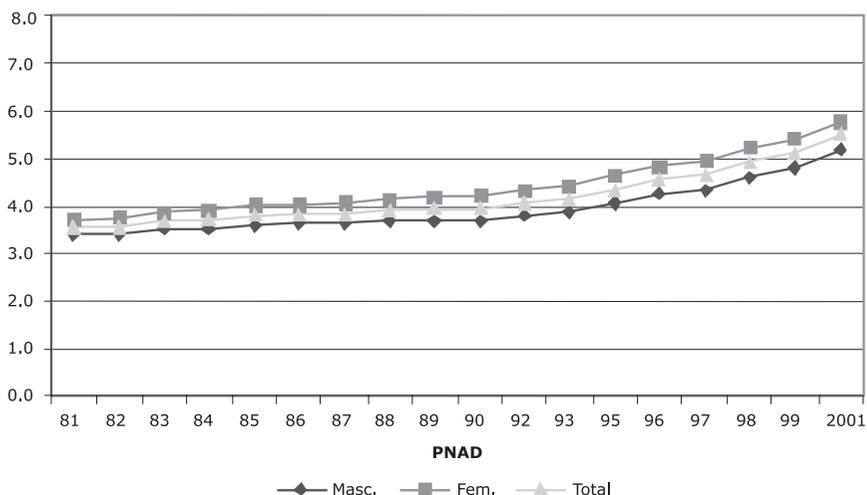
A escolaridade média é um importante indicador educacional porque na sua estimativa estão embutidas as taxas de rendimento escolar – aprovação, reprovação e evasão –, bem como o grau de atendimento do sistema de ensino. Portanto, altos níveis de atendimento escolar e taxas de aprovação tendem a elevar a escolaridade média, uma vez que há maior número de pessoas dentro da escola e que estas estão progredindo para séries mais avançadas. Por outro lado, taxas de evasão e reprovação maiores tendem a diminuir a escolaridade média. Dessa forma, esta é uma boa medida síntese das taxas de rendimento escolar e do nível de atendimento do sistema de ensino.

O Gráfico 7 mostra a evolução dessa variável para a população brasileira entre os anos de 1981 e 2001, em uma série histórica obtida pelo uso de todas as PNADs realizadas entre esses anos. Cabe lembrar que em 1991 e em 2000 as PNADs não foram realizadas porque foram anos de Censos, e que em 1994 problemas técnicos impediram a realização da pesquisa. Note-se que a escolaridade do povo brasileiro aumentou de forma lenta na década de 80 e, depois disso, passou a aumentar mais rapidamente⁶.

⁶ Um aumento na quantidade que vem acompanhado, em geral, de um incremento na qualidade. Todavia, deve-se ter em mente que nem sempre essa relação é verdadeira.

GRÁFICO 7

Anos médios de estudo da população de 7 a 25 anos total e por sexo – Brasil, 1981-2001



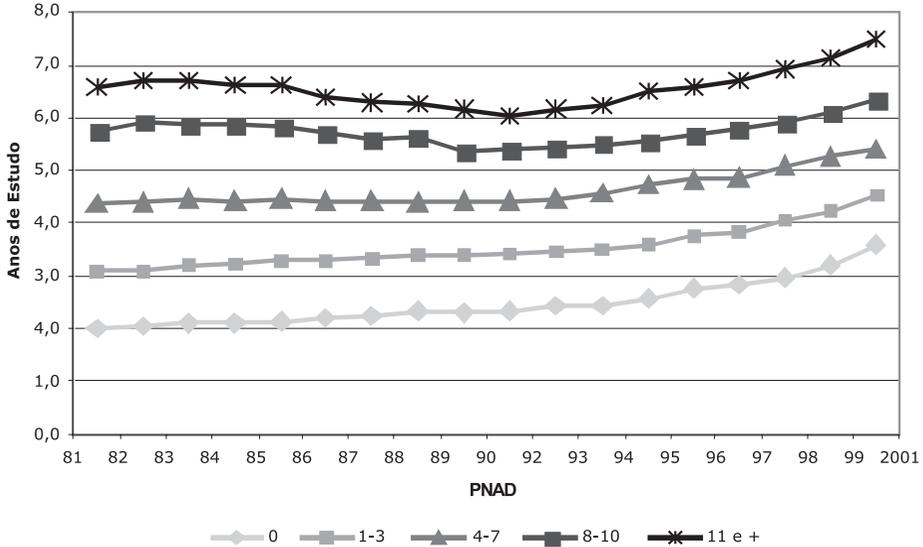
Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios 1981/2001.

Como o realizado para a taxa de analfabetismo, em muitos casos é conveniente calcular a escolaridade média para determinados grupos específicos, a fim de, entre outros aspectos, melhor identificar as camadas mais necessitadas de programas educacionais. Por exemplo, pode-se comparar a escolaridade média por faixa etária quando se deseja avaliar os avanços recentes da educação. No Gráfico 7 a escolaridade média da população de 7 a 25 anos é analisada por sexo. Como pode ser visto, ambos os sexos aumentaram sua escolaridade média no período de 1981 a 2001. Porém, as mulheres, que sempre tiveram escolaridade média superior à dos homens, apresentaram maior crescimento, ampliando a diferença entre os sexos. Tal fato pode ser explicado pela dinâmica do mercado de trabalho, uma vez que os homens tendem a nele ingressar mais cedo que as mulheres, prejudicando assim a obtenção de um melhor nível escolar.

Quando se desagrega a escolaridade média por variáveis socioeconômicas, como a educação do chefe de domicílio (Gráfico 8), percebe-se uma grande desigualdade de oportunidade educacional. Crianças que pertencem a famílias cujo chefe tem no mínimo o Ensino Médio completo (11 anos e mais de estudo) possuem escolaridade média bem maior que crianças cujo chefe do domicílio tem zero ano de estudo. Tal fato demonstra a existência de um ciclo vicioso, em que pais com baixa escolaridade transmitem pouca escolaridade aos seus filhos e assim sucessivamente. Cabe ressaltar que a escolaridade média dos três

GRÁFICO 8

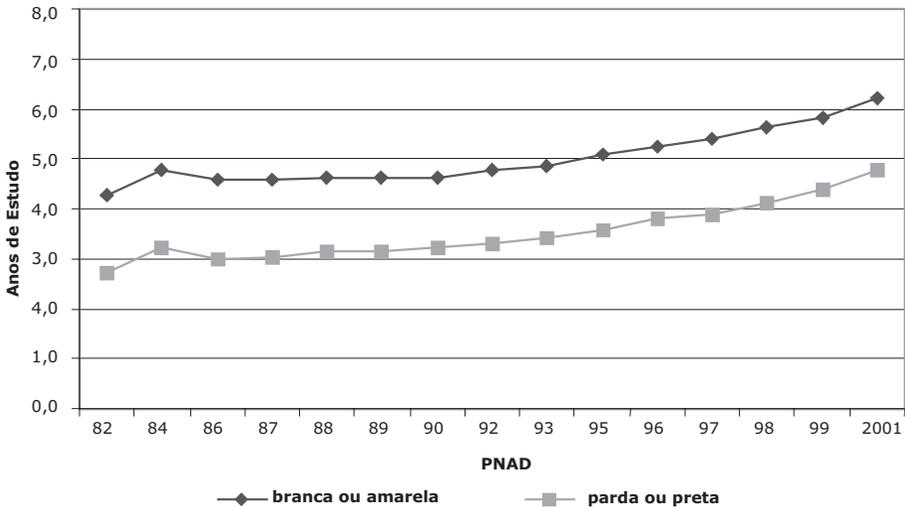
Anos médios de estudo da população de 7 a 25 anos segundo o grupo de anos de estudo do chefe de família – Brasil, 1981-2001



Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios 1981/2001.

GRÁFICO 9

Anos médios de estudo da população de 7 a 25 anos segundo a cor do chefe de família – Brasil, 1982-2001



Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios 1981/2001.

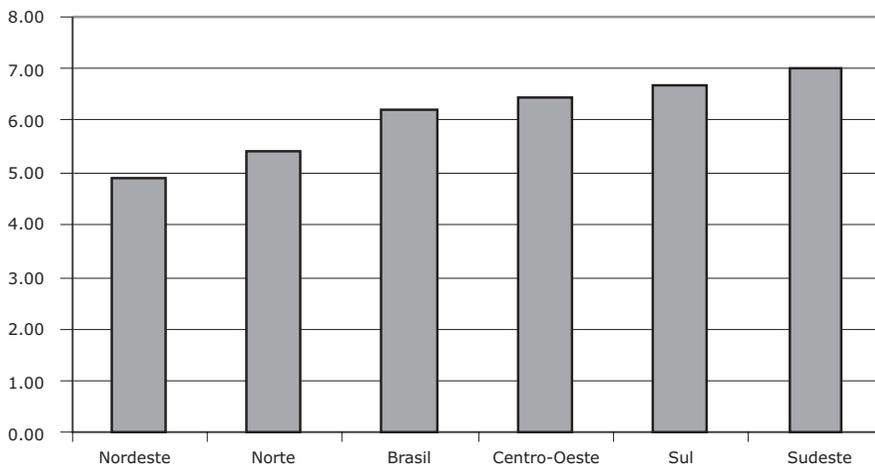
menores grupos de anos de estudo do chefe da família – as três últimas curvas do Gráfico 8 – apresentou um crescimento maior do que a dos outros dois grupos. O que indica que a diferença entre eles diminuiu nos últimos anos, ou seja, parece estar havendo uma homogeneização, mesmo que ainda fraca, nos níveis escolares de diferentes camadas sociais brasileiras.

Quando se analisa a escolaridade média por cor do chefe da família (Gráfico 9), nota-se que crianças cujas famílias são chefiadas por brancos ou amarelos possuem maior escolaridade do que aquelas nascidas em famílias com chefes pretos ou pardos⁷. Esta diferença pode estar refletindo as características socioeconômicas das famílias, já que chefes pretos e pardos são os que possuem menor educação e renda.

Quanto às diferenças regionais, percebe-se pelo Gráfico 10 que é grande a desigualdade educacional entre as regiões do Brasil. O Sudeste, o Sul e o Centro-Oeste apresentam médias de escolaridade maiores do que o Norte e o Nordeste. Em 2000, a escolaridade média nacional era de 6,23 anos de estudo para a população de 15 ou mais anos de idade, o que configura uma baixa escolaridade, pois significa que, na média, a população brasileira não tem nem o Ensino Fundamental completo. Até mesmo o Sudeste, região com maior escolaridade, com média de 7 anos de estudo, tinha um valor baixo.

GRÁFICO 10

Anos médios de estudo da população de 15 anos ou mais – Brasil e Grandes Regiões, 2000



Fonte: IBGE, Censo Demográfico de 2000.

⁷ Nos anos de 1981, 1983 e 1985 a PNAD não incluiu o quesito sobre cor ou raça, impossibilitando a desagregação da escolaridade média nesses anos.

Porcentagem da população com idade i por anos de estudo

Um indicador diretamente derivado do índice acima é o percentual da população em um determinado grupo etário por anos de estudo. O INEP normalmente o calcula para pessoas com idades entre 25 e 64 anos e entre 25 e 34 anos e agrupa os anos de estudo em cinco categorias, quais sejam: sem instrução (menos de 1 ano de estudo); Ensino Fundamental incompleto (1 a 7 anos); Ensino Fundamental completo (8 a 10 anos); Ensino Médio completo (11 a 14 anos) e Ensino Superior (15 anos ou mais de estudo). Pode-se, entretanto, utilizar outros critérios de agrupamento de anos de estudo ou outras faixas etárias, dependendo do que se pretende analisar.

A fórmula abaixo mostra como são realizados os cálculos:

$$PPAE = \frac{P_{i,c}}{P_i} \times 100,$$

onde:

$PPAE$ é a porcentagem da população por anos de estudo;

$P_{i,c}$ é a população com c anos de estudo em um determinado grupo etário i ; e

P_i é a população total no mesmo grupo etário.

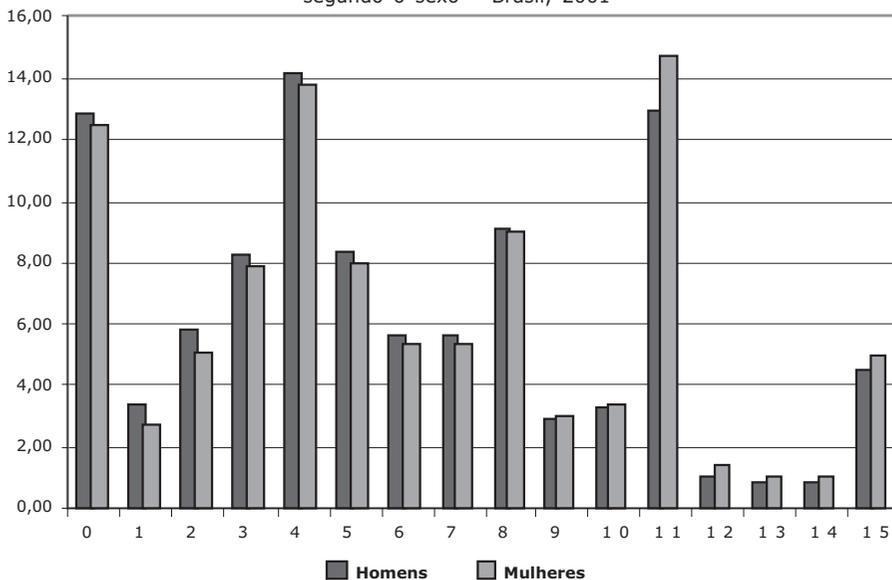
Nos Gráficos 11 e 12 são mostradas as porcentagens da população com 10 anos ou mais de idade por anos de estudo. No Gráfico 11, onde este indicador está desagregado por sexo, percebe-se nos níveis mais baixos de escolaridade uma proporção maior de homens do que mulheres, tendência que é invertida nos níveis mais elevados. Dois outros fatos chamam a atenção: a grande proporção de pessoas com zero ano de estudo (mais de 12%, próxima à taxa de analfabetismo⁸) e os picos que ocorrem para 4, 8 e 11 anos de estudo. Estes picos são devidos à maior evasão escolar na conclusão de níveis bem definidos de escolaridade: 4ª série (antigo Elementar completo), 8ª série (Ensino Fundamental completo) e 3ª série do Ensino Médio (Ensino Médio completo). Este fenômeno é chamado de efeito certificado.

Quando se analisa a porcentagem da população com c anos de estudo segundo a situação de localização do domicílio (Gráfico 12), é confirmado o maior percentual de pessoas com maior escolaridade nas áreas urbanas, relativamente às rurais. Esse fato pode ser parcialmente explicado pelas diferenças no mercado de trabalho entre esses dois setores, com uma forte presença de atividades agropecuárias na zona rural e a maior participação dos setores industriais e de serviços nas cidades.

⁸ Nem todos os que possuem zero ano de estudo são analfabetos, pois, lembrando, só se considera analfabeto aquele que não sabe ler nem escrever um bilhete simples.

GRÁFICO 11

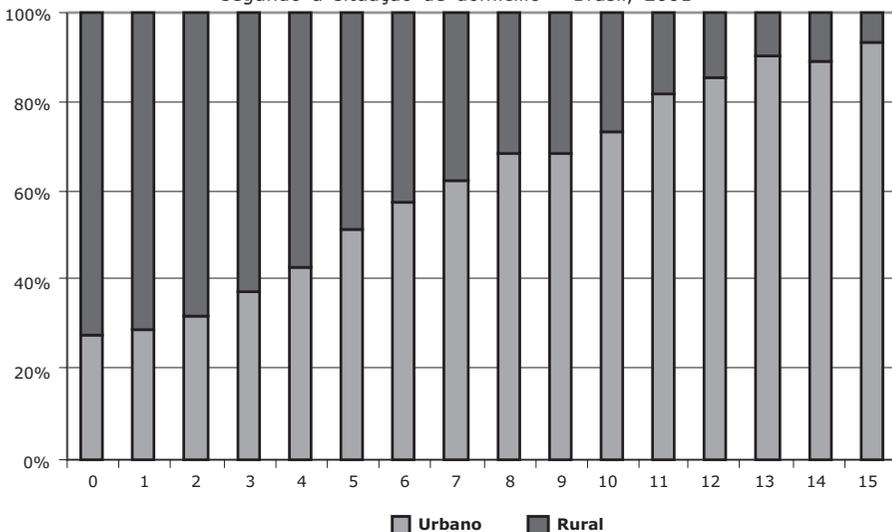
Porcentagem de pessoas com 10 anos e mais de idade com c anos de estudo, segundo o sexo - Brasil, 2001



Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios 2001.

GRÁFICO 12

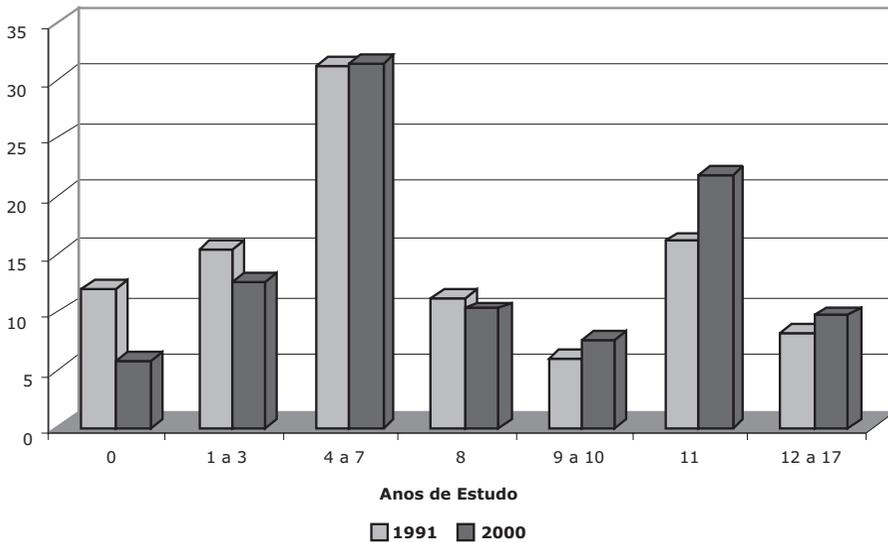
Porcentagem de pessoas com 10 anos e mais de idade com c anos de estudo, segundo a situação de domicílio - Brasil, 2001



Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios 2001.

Para comparar a evolução deste indicador entre os anos de 1991 e 2000, optou-se por agrupar os anos de estudos em: zero ano de estudo; 1 a 3 anos de estudo (analfabetos funcionais); 4 a 7 anos de estudo (Ensino Fundamental incompleto); 8 anos de estudo (Ensino Fundamental completo); 9 a 10 anos de estudo (Ensino Médio incompleto); 11 anos de estudo (Ensino Médio completo) e 12 e mais (Ensino Superior incompleto e completo). Desse modo, verificou-se uma melhora significativa na escolaridade brasileira no período, com uma diminuição da população com zero ano e com 1 a 3 anos de estudo e um aumento da porcentagem de pessoas de 25 anos com o Ensino Médio completo e o Ensino Superior incompleto ou completo.

GRÁFICO 13
Porcentagem de pessoas com 25 anos por grupos de anos de estudo
- Brasil, 1991/2000



Fonte: IBGE, Censos Demográficos de 1991 e 2000.

Taxa de atendimento escolar da população por faixa etária

Um outro indicador muito utilizado em demografia da educação é a taxa de atendimento escolar ou taxa de freqüência escolar por grupo etário. Este é um bom indicador para a avaliação de políticas recentes, dependendo da idade e do nível escolar das pessoas analisadas. Ele capta a proporção da população em uma determinada faixa etária que freqüenta a escola, podendo avaliar a

capacidade do sistema de ensino de manter as crianças e adolescentes nas escolas. Normalmente, considera-se a faixa etária adequada para se cursar determinado grau, ou seja, 4 a 6 anos para a Pré-Escola, 7 a 14 para o Ensino Fundamental, 15 a 17 para o Ensino Médio e 18 a 22-24 anos para o Ensino Superior.

A fórmula que expressa esse indicador é mostrada a seguir:

$$TAE = \frac{MAT_{(i)}}{P_i} \times 100,$$

onde:

TAE é a taxa de atendimento escolar;

MAT_i é a matrícula em todos os níveis de ensino na faixa etária selecionada; e

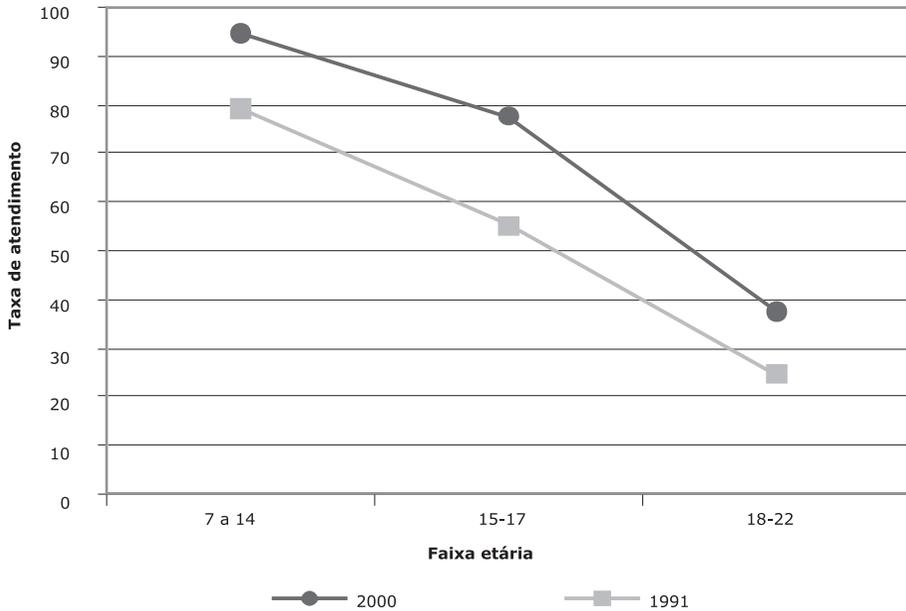
P_i é a população na mesma faixa etária.

Antes de mostrarmos alguns dos resultados obtidos para o Brasil, cabe ressaltar que esse indicador apresenta ambigüidades. Uma região pode apresentar uma taxa de atendimento escolar superior a uma outra e deter um sistema educacional muito mais precário devido a altas taxas de defasagem escolar. Portanto, com este indicador não é possível avaliar a questão da repetência, ou seja, não é possível avaliar a proporção de pessoas que se encontram cursando um nível de ensino inadequado à sua idade, o que pode ser feito através de outros tipos de indicadores, que serão vistos mais adiante.

O Gráfico 14 traz as taxas de atendimento para as faixas etárias correspondentes ao Ensino Fundamental, Médio e Superior, utilizando dados dos Censos Demográficos de 1991 e 2000. Percebe-se que entre estes dois anos ocorreu um aumento no atendimento escolar em todas as faixas etárias analisadas. Parte disso pode ser explicado pelas recentes políticas educacionais de retenção de crianças na escola, como a merenda escolar e a bolsa-escola, além da maior oferta de escolas e vagas. Para as pessoas de 7 a 14 anos, pode-se dizer que já há uma quase universalização do acesso no Ensino Fundamental, uma vez que hoje mais de 94,5% das pessoas nesta faixa freqüentam escola.

Quando se analisa a taxa de atendimento da faixa etária de 7 a 14 anos por Grandes Regiões (Gráfico 15), observa-se que em 1991 as diferenças regionais eram maiores do que em 2000. Apesar de todas as regiões apresentarem crescimento desta taxa no período analisado, a região Nordeste foi a que teve maior aumento. A diminuição das diferenças regionais pode ser atribuída, em parte, ao reflexo de políticas focalizadas para corrigi-las, dentre elas, a criação de um novo modelo de financiamento da educação – o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (FUNDEF).

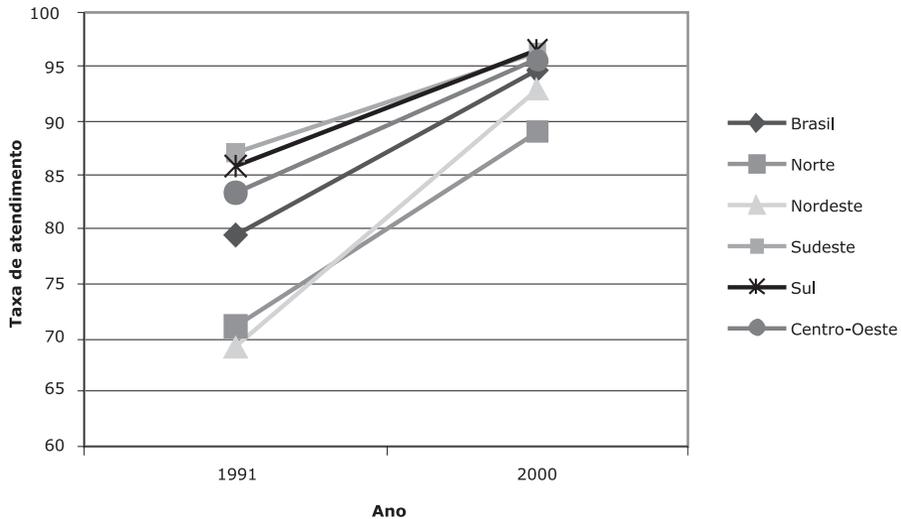
GRÁFICO 14
Taxa de atendimento por faixa etária - Brasil, 1991/2000



Fonte: IBGE, Censos Demográficos.

GRÁFICO 15

Taxa de atendimento para a faixa etária de 7 a 14 anos - Brasil e Regiões, 1991/2000

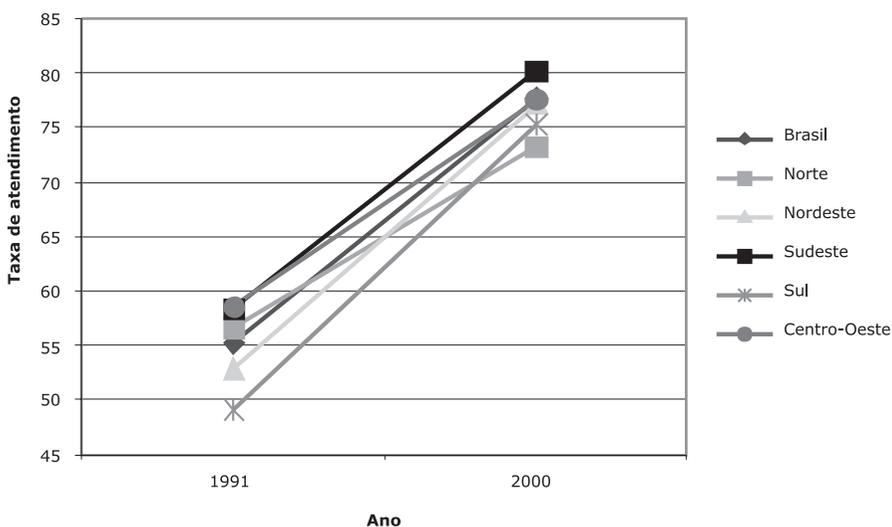


Fonte: IBGE, Censos Demográficos.

Quando se observa esta taxa para a faixa etária de 15 a 17 anos (Gráfico 16), o aspecto ambíguo fica bem evidenciado, já que a região Sul possui taxas menores que o Nordeste para os dois anos. Uma explicação pode ser a elevada repetência da região Nordeste, o que acarreta um alto percentual de pessoas ainda cursando o Ensino Fundamental. Associado a isso, como o Ensino Médio ainda não tem sido foco de políticas de retenção, há uma maior evasão das pessoas nesta faixa de idade que já concluíram o Ensino Fundamental. Dessa forma, como na região Sul a repetência é menor, os jovens de 15 a 17 anos que já concluíram o Ensino Fundamental evadem, o que causa um menor taxa de atendimento para esta faixa etária.

GRÁFICO 16

Taxa de atendimento para a faixa etária de 15 a 17 anos – Brasil e Regiões, 1991/2000



Fonte: IBGE, Censos Demográficos.

Taxa de escolarização bruta

Este indicador se parece com o anterior mas apresenta importantes diferenças metodológicas. É dado pela razão entre as matrículas em um determinado nível de ensino e a população em idade adequada de cursar tal nível. Ele possibilita avaliar o volume de matrículas nesse nível em função da demanda potencial na faixa etária adequada.

A expressão que calcula esse indicador é dada pela fórmula a seguir:

$$TEB = \frac{MAT_{(j)}}{P_i} \times 100,$$

onde:

TEB é a taxa de escolarização bruta;

MAT_j é a matrícula total em um determinado nível de ensino j ; e

P_i é a população na faixa etária adequada a esse nível de ensino.

Como o numerador desta taxa é a matrícula total independente da idade, ela pode ser inflada devido ao grande número de alunos que se encontram fora da idade adequada de cursar determinado nível de ensino. A grande proporção de pessoas fora da faixa apropriada ocorre devido à entrada tardia na escola ou à repetência.

Nos Gráficos 17 e 18 encontram-se as taxas de escolarização bruta para o Ensino Fundamental e o Médio, respectivamente. No caso do Ensino Fundamental,

GRÁFICO 17

Taxa de escolarização bruta do Ensino Fundamental – Brasil e estados, 2000

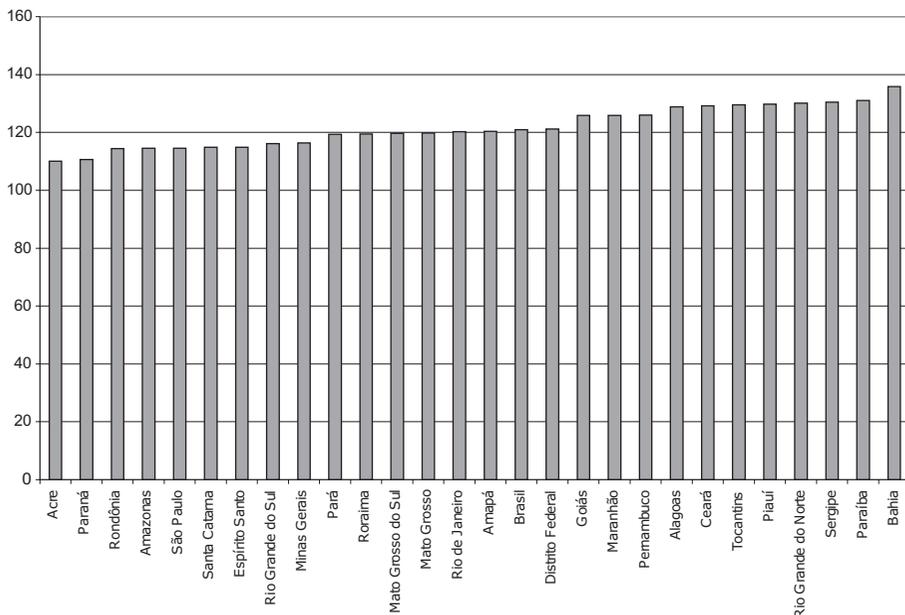
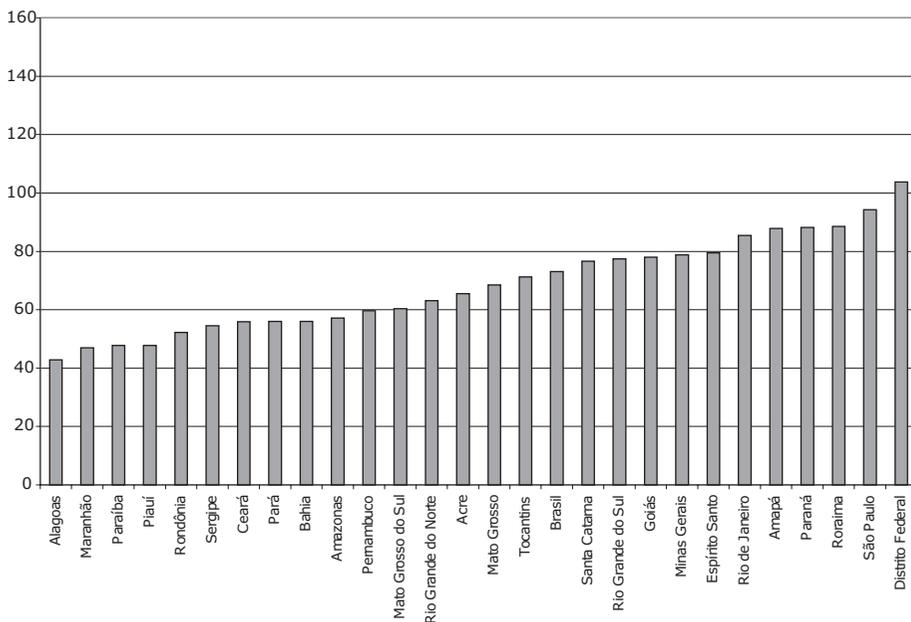


GRÁFICO 18
Taxa de escolarização bruta do Ensino Médio – Brasil e estados, 2000



Fonte: IBGE, Censo Demográfico de 2000.

a taxa de escolarização bruta ultrapassa 100% em todos os estados, evidenciando que há mais pessoas cursando este nível de ensino do que o tamanho da coorte de 7 a 14 anos. Este quadro mostra a ineficiência do sistema de ensino, com alto grau de repetência escolar.

Alguns dos estados que detêm os maiores valores para esse indicador apresentam um fraco desempenho em vários dos outros índices, como é o caso de Sergipe, da Paraíba e da Bahia. Por outro lado, no outro extremo encontramos o Acre com as menores cifras dentre todos os estados brasileiros. Percebe-se, portanto, que este indicador é um pouco ambíguo. Ele pode aumentar tanto por aspectos positivos como por negativos. Um aumento do atendimento escolar eleva a taxa de escolarização bruta, sendo este um fator positivo, já que significa que há mais crianças sendo atendidas pelo sistema de ensino. Em contrapartida, uma maior repetência pode aumentar essa taxa, o que caracteriza um fato negativo, que é a maior retenção de estudantes em determinada série.

Para o Ensino Fundamental no Brasil em 2000, observa-se que os estados com os melhores níveis escolares, como os da região Sul e São Paulo, têm uma

tendência a apresentar valores abaixo da média nacional. Esse fato indica que os aspectos negativos que trazem ambigüidade ao indicador ainda são os preponderantes.

No caso do Ensino Médio, a taxa de escolarização bruta é menor, devido, em parte, ao grande contingente de pessoas de 15 a 17 anos que ainda estão cursando o Ensino Fundamental e, em parte, à maior evasão nas idades mais velhas.

O Gráfico 18 mostra que a ambigüidade do indicador nesse nível de ensino tende para os aspectos positivos. Os estados que apresentam os maiores valores são, de forma geral, aqueles com os melhores sistemas de ensino.

Taxa de escolarização líquida

A ambigüidade observada no indicador anterior não está presente na taxa de escolarização líquida. Este indicador corresponde à razão entre as matrículas das pessoas em idade adequada para estar cursando um determinado nível e a população total na mesma idade, ou seja, indica a porcentagem da população na faixa etária que está matriculada no nível de ensino adequado. Como este indicador não capta os estudantes que estão atrasados e fora de seu nível adequado de ensino, ele é mais apropriado para avaliar a eficiência do sistema de ensino do que o anterior, já que um crescimento dessa taxa só ocorre, basicamente, por fatores positivos⁹.

A expressão que calcula esse indicador é dada pela fórmula a seguir:

$$TEL = \frac{MAT_{(ij)}}{P_i} \times 100,$$

onde:

TEL é a taxa de escolarização líquida;

MAT_{ij} é a matrícula na faixa etária adequada a um determinado nível de ensino; e

P_i é a população na mesma faixa etária.

O Gráfico 19 mostra os resultados para esse indicador para o Ensino Fundamental. Observe-se que, como esperado, e ao contrário do indicador anterior, todos os valores se encontram abaixo dos 100%. Note-se também que a ambigüidade desaparece. Os estados que reconhecidamente detêm os melhores sistemas de ensino estão entre os que têm as maiores cifras. O contrário é verdadeiro no outro extremo. Chama a atenção o desempenho de dois estados

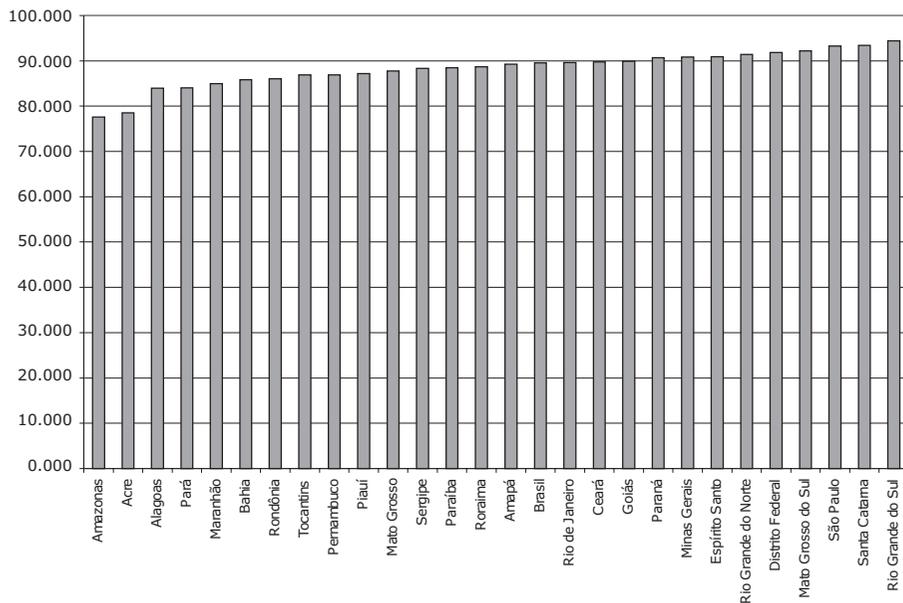
⁹ Note-se que um estudante de 10 anos de idade na 1ª série do Ensino Fundamental tem idade certa para estar no nível em que se encontra, mas não está na série adequada.

do Nordeste, Rio Grande do Norte e Ceará, que apresentam valores acima da média nacional. Esses estados mostraram um grande avanço recente em vários aspectos sociais e esse indicador está captando parte dessas melhorias.

Apesar de haver diferenças estaduais nas taxas de escolarização líquida do Ensino Fundamental, com a maioria dos estados do Sul, Sudeste e Centro-Oeste apresentando maiores taxas que os estados do Norte e Nordeste, pode-se dizer que este nível de ensino está praticamente universalizado, posto que 90% dos estudantes brasileiros de 7 a 14 anos hoje estão freqüentando o Ensino Fundamental.

GRÁFICO 19

Taxa de escolarização líquida do Ensino Fundamental – Brasil e estados, 2000

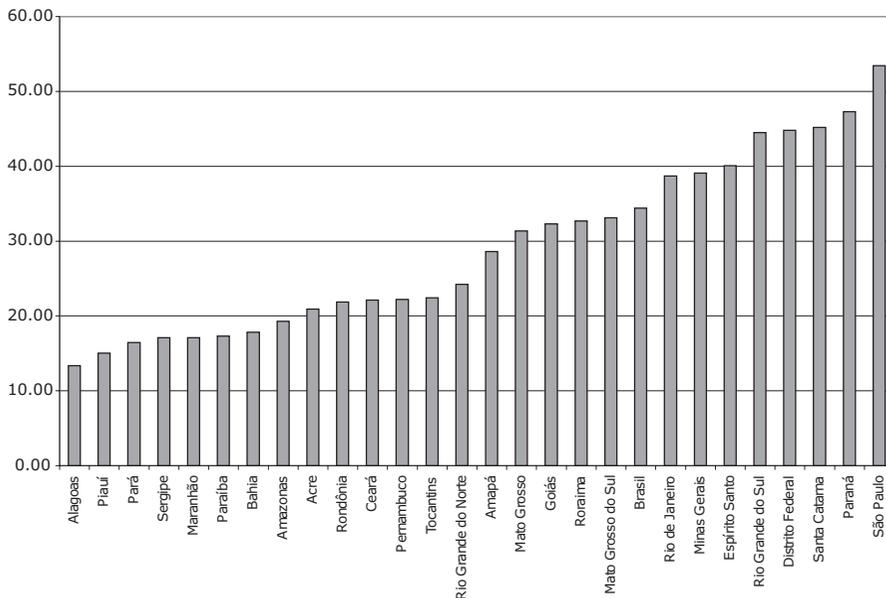


Fonte: IBGE, Censo Demográfico de 2000.

A taxa de escolarização líquida do Ensino Médio (Gráfico 20), entretanto, apresenta um nível bem abaixo do desejado: apenas 34,4% dos estudantes brasileiros de 15 a 17 anos freqüentam este nível. As disparidades entre os estados também são bem mais acentuadas, com Alagoas apresentando o pior desempenho (13,34%) e São Paulo, o melhor (53,41%).

GRÁFICO 20

Taxa de escolarização líquida do Ensino Médio – Brasil e estados, 2000



Fonte: IBGE, Censo Demográfico de 2000.

Taxa de distorção idade/série

A taxa de distorção idade/série 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, principalmente, com a alta repetência em determinada série, que é um dos principais problemas no sistema de ensino brasileiro, com graves conseqüências para os níveis de escolaridade da população.

A expressão que calcula esse indicador é dada pela fórmula a seguir:

$$TDIS_{js} = \frac{MAT_{jsi_sup}}{MAT_{js}} \times 100,$$

onde:

$TDIS_{js}$ é taxa de distorção idade/série da série s do nível de ensino j ;

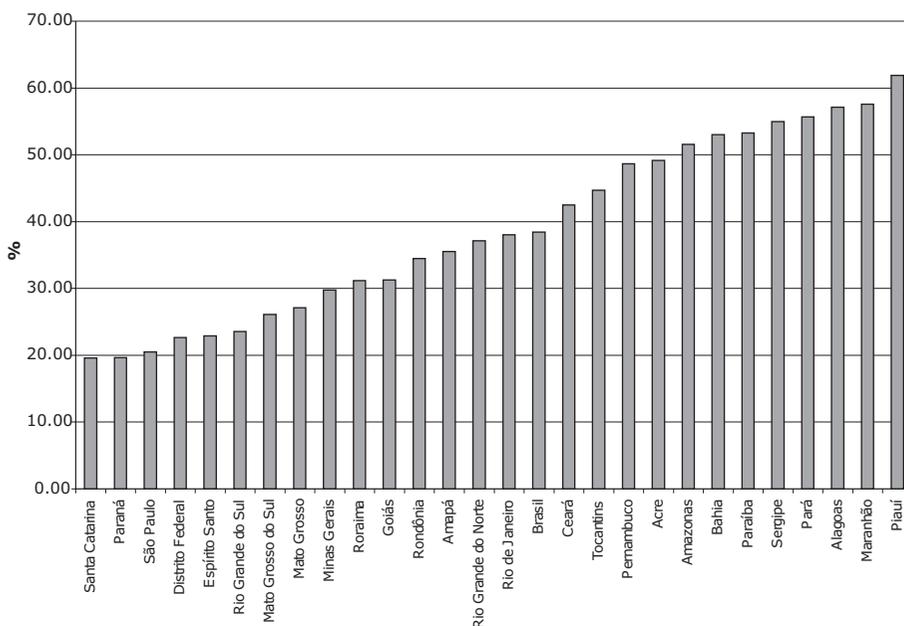
MAT_{jsi_sup} é o número de matrículas de pessoas com idade superior à idade adequada de estar cursando uma determinada série s do nível de ensino j ; e

MAT_{js} é o número total de matrículas na série s do nível de ensino j .

A metodologia de cálculo desta taxa utilizando dados do IBGE é diferente daquela proposta pelo INEP, baseada nos dados do Censo Escolar. Os Censos Demográficos e PNADs trazem informações sobre a idade e a data de nascimento da pessoa. Além disso, a data de referência dessas bases, como já comentado, é no segundo semestre do ano. Esses dois aspectos permitem determinar com clareza quais estudantes têm idade adequada para estar cursando uma determinada série e quais não têm. Por outro lado, o Censo Escolar não conta com a informação sobre a idade do indivíduo, nem sobre o mês de nascimento do estudante. A informação presente nessa base é somente o ano de nascimento da pessoa. Além disso, a data de referência da base é maio, o que impossibilita o uso da mesma metodologia aplicada na base do IBGE. O indicador distorção idade/série, quando calculado pelo Censo Escolar, inclui um maior grupo de estudantes com idade adequada.

No Gráfico 21 estão representadas as taxas de distorção idade/série da 1ª série do Ensino Fundamental para o Brasil e estados em 2000. Escolheu-se essa

GRÁFICO 21
Taxa de distorção idade/série para a 1ª série do Ensino Fundamental
- Brasil e estados, 2000



série por ser uma das que registram maior repetência. Como pode ser observado, em 2001 havia no país quase 38,5% de alunos matriculados na 1ª série que estavam fora da idade adequada. Os estados do Nordeste são, em geral, os que apresentam as maiores taxas, sendo o Piauí o pior estado, com quase 70% das matrículas referentes a pessoas em idade inadequada. A distorção idade/série, causada principalmente pelas altas taxas de repetência, é um dos principais problemas brasileiros, prejudicando o financiamento do sistema de educação, visto que os alunos demoram mais de um ano para concluir uma série.

Razão professor por mil habitantes

Os dois próximos indicadores a serem discutidos relacionam o número de professores com a população em geral e com o número de alunos. Nas pesquisas domiciliares é possível determinar o número de professores em determinada unidade geográfica, por meio do quesito referente à ocupação do trabalho principal¹⁰. Desta forma, é possível calcular indicadores relacionados à disponibilidade de professores, como a razão professor por mil habitantes.

A expressão que calcula esse indicador é dada pela fórmula a seguir:

$$\text{prof/hab} = \frac{NP}{P} * 1000,$$

onde:

prof/hab é a razão professor por mil habitantes;

NP é o número de professores; e

P é a população total.

Esta razão pode também ser calculada utilizando-se como população de referência aquela em idade escolar, ou pode ser ainda mais refinada, calculada para cada nível de ensino. Nesse caso, considera-se o número de professores em um determinado nível de ensino sobre a população em idade adequada de cursar esse nível.

Este indicador, como já foi ressaltado, é importante na medida em que indica a disponibilidade da oferta de professores para as unidades geográficas, apontando regiões onde há escassez desses profissionais. Desse modo, pode contribuir para a formulação de políticas que minimizem tais restrições de oferta.

¹⁰ Todavia, deve-se levar em consideração que por meio deste quesito só são captados os professores que têm esta profissão como a principal.

Razão aluno/professor

A razão aluno/professor também é um indicador de recursos educacionais. Ela é muito utilizada como um indicador de qualidade da educação. A hipótese é de que, em classes pequenas, os alunos têm maior atenção e, por conseqüência, um melhor desempenho. Existem, porém, controvérsias com relação a esta hipótese, discutidas na literatura nacional e internacional sobre o assunto (ver Hanushek, 2003, e Krueger, 2003).

A expressão que calcula esse indicador é dada pela fórmula a seguir:

$$\text{Aluno/prof} = \frac{M_j}{NP_j},$$

onde:

Aluno/prof é a razão aluno/professor;

M_j é a matrícula no mesmo nível de ensino; e

NP_j é o número de professores em um determinado nível de ensino.

Também a título de ilustração, apresentamos a seguir alguns dados comparativos entre países (Tabela 3). Todos os países com renda mais elevada têm uma razão aluno/professor abaixo de 20, indicando que os estudantes dessas regiões têm um melhor acompanhamento. Os países de renda intermediária apresentam valores um pouco mais elevados, entre 20 e 30 alunos por professor. A exceção é a Hungria, que apresenta valores muito mais baixos, possivelmente pela herança de seu recente passado comunista. As regiões com menor renda

TABELA 3
Número de alunos por professor para alguns países selecionados

Grupo de renda	País	Número de alunos por professor (ou ao contrário)
Países com elevada renda <i>per capita</i>	Alemanha	18
	Estados Unidos	16
	França	19
	Japão	18
Países com renda <i>per capita</i> intermediária	Brasil	23
	Chile	27
	Hungria	11
	México	29
Países com baixa renda <i>per capita</i>	Chade	62
	China	24
	Índia	63
	Mali	66

por habitante mostram valores muito mais elevados desse índice, sinalizando a precariedade do ensino nessas áreas. A China é o único país que destoa desse quadro geral, pois apresenta um valor próximo dos países de renda intermediária, apesar de ter uma renda média mais baixa. Além de seguir políticas diferenciadas no campo socioeconômico, por ser teoricamente comunista, a China está entre os países que mais crescem economicamente, e esse investimento na educação deve ter uma relação direta com isso.

Conclusão

Este capítulo buscou apresentar algumas características das bases de dados domiciliares do IBGE, os Censos Demográficos e as PNADs, principalmente no que tange às variáveis educacionais. Essas bases são bastante ricas e possibilitam o estudo da educação de diversas perspectivas, incluindo diferenciações quanto a aspectos demográficos e socioeconômicos, como nível de renda, educação dos pais, sexo, grupo ocupacional etc., por meio da utilização de uma ampla gama de indicadores. Num segundo momento foram abordados alguns desses indicadores que permitem, quando aplicados sistematicamente, fazer uma avaliação muito completa da evolução da educação do Brasil.

Uma das conclusões possíveis a respeito dos dados apresentados é a de que o Brasil apresentou uma marcante melhora em todos os indicadores, sobretudo ao longo da década de 90. Grande parte da melhoria nos indicadores ocorreu de maneira mais marcante no Ensino Fundamental do que no Ensino Médio, indicando que o grande desafio futuro está relacionado a este último nível de ensino, tanto em termos do atendimento para a faixa etária de 15 a 17 anos, quanto com vistas a superar o atraso escolar destes jovens.

Também se verificou neste período uma diminuição das diferenças regionais, embora elas ainda persistam, com os estados do Norte e Nordeste apresentando os piores indicadores. Tal situação ratifica a necessidade de políticas voltadas para as áreas mais carentes em termos educacionais.

Todavia, deve-se ter em mente que o ensino brasileiro, apesar dos recentes avanços, que de forma alguma são desprezíveis, ainda é muito precário, mesmo quando comparado ao de países de renda por habitante igual ou semelhante à nossa. Esse fato pode ser verificado em uma série de análises sobre o assunto publicadas em Birdsall e Sabot (1996), livro que discute a questão da baixa qualidade do ensino no Brasil e deveria ser consultado em estudos mais avançados sobre o tema da educação.

O baixo desempenho educacional dos brasileiros acarreta ao país vários problemas econômicos e sociais. Pelo lado econômico, prejudica o crescimento da economia e aumenta o atraso tecnológico do país em relação às outras economias. 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. Os trabalhadores não-qualificados possuem um baixo rendimento em relação aos qualificados e com isso investem pouco na educação dos filhos, que, possivelmente, também terão um baixo rendimento, alimentando, desta forma, um ciclo vicioso.

Em face do apresentado acima, é de suma importância melhorar o sistema de ensino brasileiro. O primeiro passo para isso é garantir um conhecimento detalhado sobre a educação brasileira e sua evolução recente em diferentes recortes temporais e geográficos. As bases de dados e indicadores aqui discutidos constituem uma ferramenta fundamental para tal empreendimento.

Referências bibliográficas

BIRDSALL, N. e SABOT, R. (orgs.). **Opportunity foregone** – education in Brazil. Washington, D.C.: Inter-American Development Bank, 1996.

IBGE. **Censo Demográfico de 1991**: documentação dos microdados da amostra. Rio de Janeiro: IBGE, 1996.

_____. **Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2001**: notas metodológicas. Rio de Janeiro: IBGE, 2002.

_____. **Censo Demográfico de 2000**: características gerais da população: resultados da amostra. Rio de Janeiro: IBGE, 2003. Disponível em: <www.ibge.gov.br>. Acesso em: out. 2003.

_____. **Censo Demográfico de 2000**: documentação dos microdados da amostra. Rio de Janeiro: IBGE, 2003. CD. 2ª edição.

_____. **Censo Demográfico de 2000**: resultado do universo. Rio de Janeiro: IBGE, 2003. Disponível em: <www.ibge.gov.br>. Acesso em: out. 2003.

HANUSHEK, Eric A. (2003). The failure of input-based schooling policies. **The Economic Journal**, 113 (February), F64-F98, Royal Economic Society.

JANNUZI, P. M. **Indicadores sociais no Brasil** – conceitos, fontes de dados e aplicações. Campinas/São Paulo: Editora Alínea, 2001.

KRUEGER, Alan B. (2003). Economic considerations and class size. **The Economic Journal**, 113 (February), F34-F63, Royal Economic Society.

SEEC/INEP/MEC. **Geografia da educação brasileira**. Brasília: INEP, 2002.

SOARES, S. e LIMA, A. F. A mensuração da educação nas PNADs da década de 1990. **Textos para Discussão**, nº 928, IPEA, Rio de Janeiro, 2002.

