

Medidas de nível socioeconômico em pesquisas sociais: uma aplicação aos dados de uma pesquisa educacional¹

Maria Teresa Gonzaga Alves

Instituto de Ciências Humanas e Sociais
Universidade Federal de Ouro Preto

José Francisco Soares

Universidade Federal de Minas Gerais
Departamento de Ciências aplicadas à Educação
GAME - Grupo de Avaliação e Medidas Educacionais

Resumo: Em qualquer sociedade, diferenças entre os indivíduos são consistentemente observadas quanto ao lugar que ocupam na hierarquia social. Tais diferenças se associam profundamente às oportunidades educacionais, às trajetórias ocupacionais, ao acesso aos bens e serviços, ao prestígio social, ao comportamento político e social, etc. Neste artigo, assume-se que todas essas diferenças podem ser descritas adequadamente por um único construto denominado de nível socioeconômico (NSE). Este trabalho tem como objetivos principais: discutir formas de medir o NSE nas pesquisas sociais e produzir uma medida de NSE que expresse as variações entre famílias representadas nessas pesquisas, bem como interpretar a escala proposta. A consistência da proposta teórica é verificada com dados da Pesquisa Geração Escolar 2005 - Polo Belo Horizonte.

Palavras-chave: nível socioeconômico; escalas e medidas; metodologia; pesquisa educacional

Abstrac: In any society, individuals differ markedly in relation to their place in the social hierarchy. Those differences are deeply associated to each individual educational opportunities, occupational history, access to goods and services, social prestige and political and social options. In this paper, we assume that all those differences can be adequately described by a unique construct named SES – Socioeconomic Status. Given this, the main purposes of this paper is to discuss different forms of measuring the SES in applied social science research and to introduce a specific measure, with the appropriate interpretation, for this construct. The usefulness of the new measure is verified with the data from a longitudinal educational survey named GERES 2005 – acronym for School Generation 2005.

Keywords: socioeconomic status; scaling; social science methodology; educational research

¹ Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig), que viabilizou a produção dos dados utilizados neste trabalho por meio de apoio à pesquisa GERES (Processo No. 847/05).

Introdução²

Em qualquer sociedade, diferenças entre os indivíduos são consistentemente observadas quanto ao lugar que eles ocupam na hierarquia social. Reconhece-se que tais diferenças se associam às oportunidades educacionais, às trajetórias ocupacionais, ao prestígio social, ao acesso aos bens e serviços, ao comportamento político e social etc. O estudo dessas diferenças, seja como um fenômeno a ser explicado ou sua associação a outros fenômenos sociais, constitui uma área de grande importância nas pesquisas sociais. Mas para que as diferenças na hierarquia social possam ser apreendidas na pesquisa empírica é necessário um esforço no sentido de definir e operacionalizar medidas dessas diferenças.

Uma alternativa é a distinção dos indivíduos por meio de variáveis diretamente observadas, tais como os níveis de escolaridade ou de renda. Outra forma é a estratificação dos indivíduos ou das famílias por meio de uma medida sintética de posição social. Essa é a opção deste artigo, no qual a medida utilizada recebe o nome de nível socioeconômico - NSE. Do ponto de vista mais operacional, o NSE é tomado como um construto teórico, ou seja, uma variável latente (não diretamente observada) cuja medida é feita pela agregação de informações sobre: a educação, a ocupação e a riqueza ou rendimento dos indivíduos.

O nível socioeconômico aparece em inúmeros estudos como variável explicativa ou de controle para a análise de diversos fenômenos sociais. No entanto, não há um consenso na literatura sobre sua conceituação, bem como sobre como medi-lo nas pesquisas empíricas. Há vários aspectos relacionados ao NSE que vêm merecendo debate entre os cientistas sociais. Por exemplo, a sua base de conceituação teórica, o tipo de medida - se contínua ou categórica -, os fatores a serem considerados na produção da medida - se variáveis isoladas ou em forma de um índice -, a definição de um esquema de classificação das variáveis consideradas e a importância relativa do pai e da mãe para a definição do NSE das famílias (BOLLEN et al, 2001; BUCHMANN, 2002; CIRINO et al, 2002; ENSMINGER et al, 2007; ENTWISLE e ASTORE, 1994; ERIKSON, 1984; ERIKSON e GOLDTHORPE, 1992a e 1992b; HAUSER e WARREN, 1997; LYNCH e KAPLAN, 2000; MAY, 2006; MUELLER e PARCEL, 1981; OAKES e ROSSI, 2003; OSBORN, 1987; SIRIN, 2005; WHITE, 1982).

² A primeira autora agradece à Neuma Aguiar, professora titular do Departamento de Sociologia da Universidade Federal de Minas Gerais, pela oportunidade de participar das discussões teóricas e metodológicas sobre a construção de índices para medidas de nível socioeconômico nas pesquisas de estratificação e mobilidade social durante o estágio de pós-doutorado em 2006/2007. Essa experiência em muito estimulou e contribuiu para o desafio de produção desse artigo. Naturalmente, as opções metodológicas deste artigo são de inteira responsabilidade dos autores e não coincidem necessariamente com as posições da referida professora.

O principal objetivo deste trabalho é apresentar uma metodologia para medir o NSE, na acepção introduzida acima, através da Teoria da Resposta ao Item, uma técnica de agregação de indicadores de uma variável latente. Como referência empírica, serão utilizados os dados de um survey educacional, a pesquisa GERES - Pesquisa Longitudinal da Geração Escolar 2005³, cujo questionário contextual fornece as informações necessárias para a implementação da metodologia proposta.

O trabalho está organizado em quatro seções, além desta introdução e da conclusão. Na próxima seção, faz-se uma breve revisão das teorias sociológicas que fundamentam a definição do construto NSE assumida pelos autores. Na segunda seção, são discutidos aspectos relacionados aos fatores considerados para a construção das medidas de NSE. Na terceira seção, são descritos os dados e os procedimentos metodológicos utilizados na preparação desses dados e na produção da escala de NSE. Os resultados são descritos na quarta seção. Na conclusão, discutem-se as possibilidades de generalização dos resultados para outras pesquisas.

Revisão da Literatura

A sociologia tem uma longa tradição metodológica em produzir medidas correlatas ao nível socioeconômico no escopo das pesquisas sobre estratificação e mobilidade social. Na maioria desses estudos, os postos ocupacionais que estruturam o sistema produtivo e o mercado de trabalho constituem a base conceitual e operacional para a definição desse tipo de medida. Para isso, nos estudos empíricos, parte-se do registro de todas as ocupações de uma sociedade, por exemplo, as que aparecem nos censos demográficos. Em seguida, essas ocupações são codificadas em medidas mais manejáveis e sociologicamente relevantes de acordo com as preferências e questões substantivas do pesquisador, considerando, por exemplo, suas características quanto às credenciais exigidas, o grau de especialização e responsabilidades envolvidas, bem como o retorno financeiro.

³ A pesquisa GERES é um estudo longitudinal que acompanha uma coorte de alunos (cerca de vinte mil) a partir do primeiro ano de escolarização no Ensino Fundamental (alunos com sete anos de idade), em mais de trezentas escolas (públicas e privadas) de cinco cidades brasileiras. A pesquisa está sendo realizada com a parceria da Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio (Laboratório de Avaliação da Educação), da Universidade de Campinas – Unicamp (Laboratório de Observação e Estudos Descritivos), da Universidade Estadual de Mato Grosso do Sul (UEMS), da Universidade Federal da Bahia – UFBA (Centro de Estudos Interdisciplinares para o Setor Público – ISP) e da Universidade Federal de Minas Gerais – UFMG (Grupo de Avaliação e Medidas Educacionais – GAME). Detalhes sobre o GERES podem ser consultados no sítio <www.geres.ufmg.br> (consultado em fevereiro de 2008).

Exemplos bem conhecidos de trabalhos com orientação empírica aplicados para a análise dos países capitalistas contemporâneos são os esquemas de classe social para a análise de dados amostrais ou censitários elaborados pelos sociólogos Erik Olin Wright e John H. Goldthorpe. Wright (1985) divide o sistema ocupacional em doze categorias que tipificam as classes sociais. Ele toma como ponto de partida a divisão clássica marxista do modo de produção capitalista entre a classe da burguesia e a do proletariado, mas incorpora as posições intermediárias, de acordo com as credenciais e a posição dos trabalhadores na organização das sociedades atuais. A tipologia de classe definida por Goldthorpe (1987) contém elementos muito semelhantes aos de Wright. Porém, ele incorpora a dimensão do mercado para definir as barreiras entre as classes, o que o aproxima da perspectiva weberiana para a definição de classe econômica. No esquema proposto pelo autor, são reconhecidas sete categorias de classes, que resultam da agregação das ocupações segundo a situação de mercado (níveis salariais, seguridade econômica e chances de melhoria econômica) e a situação de trabalho (autoridade e controle da produção).

Para o propósito deste trabalho, importa destacar que nesse tipo de análise as medidas empíricas de classe social são categóricas, no sentido que os autores definem as posições de classes em termos de posições antagônicas. Elas não são entendidas como em uma hierarquia de classes, ou seja, os autores não propõem uma ordenação das classes entre mais altas ou mais baixas, embora essa leitura seja parcialmente possível (SCALON, 1998).

A noção de nível socioeconômico tem uma perspectiva diferente ao enfatizar as medidas contínuas para descrever a estrutura social. Essa é reconhecida como aplicação empírica da teoria funcionalista, que explica o fenômeno da estratificação social pelas diferenças na realização de status (status achievement) dos indivíduos e reconhece uma estrutura hierárquica da sociedade com base nas ocupações (BLAU e DUNCAN, 1967). O estudo mais influente é o trabalho de Duncan (1963), que desenvolveu uma escala de status socioeconômico dos títulos ocupacionais registrados no censo norte-americano de 1950. Nesse estudo, o status socioeconômico era captado por um indicador de prestígio das diferentes ocupações. A escala elaborada por Duncan tomou como ponto de partida uma hierarquização do prestígio ocupacional obtida em um survey externo ao Censo conduzido alguns anos antes. Como esse survey não mediu o prestígio de todas as ocupações registradas no censo, Duncan propôs um índice associando o ranking de prestígio ocupacional com as características dessas posições em relação à educação e renda registradas no Censo. A hipótese subjacente é que a qualificação para a ocupação se dá por meio da educação, e a remuneração é consequência da ocupação, ou seja, a educação e a renda são, respectivamente, causa e efeito do status ocupacional. Com o uso de técnica de regressão múltipla, Duncan pôde

derivar um índice socioeconômico para todas as ocupações. Esse índice foi utilizado em inúmeros estudos norte-americanos e sofreu algumas atualizações para a inclusão de ocupações mais recentes (por exemplo, NAKAO e TREAS, 1994; HAUSER e WARREN, 1997), bem como inspirou estudos em vários países do mundo (TREIMAN, 1977)⁴.

No entanto, como argumentam Ganzeboom, De Graaf e Traiman (1992), considerando a complexidade da estrutura ocupacional das sociedades capitalistas contemporâneas, não é um procedimento tão simples generalizar um índice de NSE a partir de escores de prestígio ocupacional. Pesquisas sobre prestígio ocupacional são viáveis apenas com um número limitado de ocupações. Dessa forma, na maioria dos países, as escalas de NSE são produzidas a partir de uma média ponderada de características dos indivíduos que exercem as ocupações, geralmente a educação e renda, às vezes também considerando a idade ou o sexo. Os procedimentos metodológicos variam entre os pesquisadores, mas os resultados são muito parecidos, geralmente revelando que a educação tem um peso um pouco maior do que a renda na definição do índice de NSE, e outros fatores têm peso bem menor. Tendo isso como pressuposto, Ganzeboom, De Graaf e Traiman (1992) propuseram uma medida de NSE que independe de escores de prestígio ocupacional com o objetivo de ser aplicado em pesquisas internacionais. Eles construíram uma escala a partir das informações sobre educação, ocupação e renda dos indivíduos disponíveis em surveys ou censos nacionais à qual denominaram de *International Socio-economic Index* (ISEI). O objetivo dos autores é que o ISEI possa ser aplicado para a atribuição de um valor de NSE em qualquer estudo que tenha o registro da ocupação dos indivíduos pesquisados, sem a necessidade de coletar dados primários sobre educação e a renda. Na escala do ISEI, o menor escore, 16, foi atribuído às ocupações manuais, tais como serventes, empregados domésticos e trabalhadores rurais, e o maior, 90, à ocupação de juiz⁵. O ISEI tem ampla aceitação nos estudos empíricos no campo da educação, por exemplo, na pesquisa comparativa internacional de desempenho escolar PISA⁶.

⁴ Treiman (1977) comparou 85 estudos de prestígio ocupacional em 60 diferentes países. Ele concluiu que as hierarquias de prestígio quase não variavam no espaço e no tempo. A correlação entre os escores obtidos em cada estudo com uma escala padrão construída apresentou valores de 0,68 a 0,97, com uma correlação média de 0,91.

⁵ O ISEI foi desenvolvido a partir da Classificação Uniformizada das Ocupações (CISCO68, cuja sigla em inglês é ISCO68). A escala foi criada combinando-se dados comparáveis sobre educação, renda e ocupações masculinas de 16 países. O ISEI foi revisto em outro estudo (GANZEBOOM e TREIMAN, 1996) para incluir a última revisão publicada do CISCO (ISCO1988). A escala do ISEI foi validada quanto à sua aplicabilidade em diferentes contextos e possui forte correlação com outras escalas, como a escala de prestígio conhecida como SIOPS (sigla em inglês para Standard International Occupational Prestige Scale) produzida por Treiman (1977).

⁶ Entretanto, no PISA, o ISEI foi utilizado apenas como um dos componentes do índice denominado Index of economic, social and cultural status, conforme descrito no relatório oficial: PISA-OECD Knowledge and Skills for Life: First Result from PISA 2000. p 221, 2001.

A opção metodológica para a construção do ISEI é compatível com vários outros estudos. Meta-análises com resultados de pesquisas sobre estratificação educacional mostram que a ocupação, a educação e a renda dos indivíduos ou dos responsáveis pelas famílias, como variáveis independentes ou na forma de índice, são as variáveis mais usadas nos estudos correlacionais da área (BUCHMANN, 2002; CIRINO et al, 2002; SIRIN, 2005; WHITE, 1992).

No Brasil, pesquisas em estratificação social e mobilidade têm estimulado a produção de esquemas de classificação socioeconômica adaptados à realidade do país. Os estudos nacionais mais influentes utilizam os dados produzidos pelo Instituto Brasileiro de Geografia e Estatística (IBGE) no âmbito da Pesquisa Nacional por Amostra de Domicílio (PNAD) e do Censo Demográfico. O esquema de classificação socioeconômica elaborado por Pastore (1979) com base nos dados da PNAD 1973, atualizado por Pastore e Valle Silva (2001) com os dados da PNAD 1996, pode ser descrito como um esquema hierárquico de classificação socioeconômica. Esses estudos, que tiveram como objetivo a análise da mobilidade social no Brasil em duas décadas distintas, derivaram uma escala de status socioeconômico combinando o nível educacional e o nível de rendimentos dos indivíduos dentro de cada título ocupacional registrados na PNAD. A escala obtida foi então dividida em seis estratos hierarquizados: (1) baixo-inferior; (2) baixo-superior; (3) médio-inferior; (4) médio-médio; (5) médio-superior e (6) alto.

Há outras referências nacionais dignas de nota, como os trabalhos de Scalon (1998) e Santos (2002, 2005). Com perspectivas teóricas diferentes, mas também com base nos dados da PNAD, esses estudos propõem alternativas de classificação socioeconômica que consideram não as hierarquias de status, mas a situação das classes sociais a partir da análise das posições das ocupações na produção e no mercado de trabalho. As classes assim descritas e organizadas em categorias ocupam posições distintas, não diretamente hierarquizáveis.

Em todos esses estudos há um consenso quanto à importância da ocupação dos indivíduos na definição de sua posição social. A ocupação é considerada um indicador singular para a análise da desigualdade social porque pode informar, ao mesmo tempo, os requisitos educacionais e o retorno econômico para cada posto ocupacional (HAUSER e WARREN, 1997). No entanto, se reconhece também que nem sempre é possível descrever os parâmetros básicos da estrutura de estratificação social apenas com o conhecimento dos títulos ocupacionais. Sabe-se que as categorias ocupacionais possuem uma enorme variabilidade em termos de credenciais e retorno financeiro (PASTORE VALLE SILVA, 2002, p.20). Por esse motivo, a definição da estrutura social através de índices de NSE construídos a partir da agregação de um número distinto de dimensões é uma opção analítica adotada por muitos pesquisadores. Além disso, na prática da pesquisa empírica, a definição de uma medida de NSE depende não só das opções teóricas do

pesquisador, mas também da disponibilidade de dados adequados para esse tipo de análise.

Medida do Nível Socioeconômico

Neste trabalho, como dito na Introdução, assume-se que o NSE das famílias é um construto teórico, que não pode ser diretamente medido, mas que se manifesta na educação, ocupação e renda dos responsáveis pela família. Nesta seção, são descritos os procedimentos para a coleta desses dados nas pesquisas sociais e as opções metodológicas deste trabalho.

A informação sobre a educação é a mais usual em questionários e a mais facilmente coletada. Alguns pesquisadores preferem operacionalizá-la pela contagem dos anos de estudos, como uma variável discreta, outros preferem categorizá-la segundo os níveis educacionais, reconhecendo nesses as barreiras estruturais que caracterizam os sistemas de ensino. Além disso, essa variável surge em alguns estudos como indicador único do NSE, devido à forte correlação entre rendimento e anos de estudo. Porém, com a expansão generalizada do ensino verificada nos últimos anos e diversos incentivos para o acesso aos níveis educacionais mais altos para as camadas antes excluídas da escola, é razoável supor uma redução do efeito discriminador dos anos de estudos para as gerações mais jovens. De fato, as evidências empíricas com dados de avaliação educacional no Brasil mostram que a estratificação educacional continua muito forte no país, porém, ela vem se manifestando cada vez mais e principalmente por meio da desigualdade ao acesso ao ensino de qualidade e não apenas pelo acesso à educação (FERRÃO-BARBOSA e FERNANDES, 2001; SOARES, 2005). Por esse motivo, argumenta-se que a educação, de forma isolada, constitui um indicador limitado para exprimir o NSE dos indivíduos. Esse exige o conhecimento desta informação, mas o NSE não será medido apenas por esta dimensão.

A preocupação com a coleta de dados sobre a renda dos indivíduos e das famílias também é usual nos questionários das pesquisas sociais, mas a forma de medi-lo está longe de ser padronizada entre os estudos. Porém, obter respostas fidedignas para uma pergunta sobre a renda nem sempre é possível, mesmo quando apresentada em forma de categorias (faixas de rendimento). Se a pergunta for respondida por membros da família que não são seus responsáveis principais, a frequência de respostas inválidas costuma ser bem elevada. Mas, mesmo entre os responsáveis, há muitos que não respondem a esta questão, por exemplo, aqueles que têm uma inserção precária no mercado de trabalho, os que têm um rendimento muito variado, ou mesmo os que preferem não revelar a renda. Além disso, os economistas alertam para a distinção entre a renda do trabalho e a renda do capital, sendo que os detentores de mais capital tendem muito mais a omitir a

própria renda. Nos surveys educacionais, frequentemente respondidos pelos próprios alunos, esse tipo de pergunta dificilmente faz parte dos questionários.

Uma alternativa à pergunta direta sobre a renda dos indivíduos é a coleta de dados que possibilitem inferir indiretamente a riqueza da família. Nos questionários dos surveys é comum a inclusão de vários itens sobre consumo de itens de conforto e contratação de empregados domésticos para essa finalidade. A medida do NSE pelo consumo de bens e serviços busca reconstruir as hierarquias sociais conhecendo a destinação que as famílias ou os indivíduos dão à riqueza obtida devido a sua origem social. A agregação desses itens juntamente com a informação sobre a escolaridade do responsável pela família constitui a base para o cálculo do Critério de Classificação Econômica Brasil, proposto pela ANEP - Associação Nacional de Empresas de Pesquisa - que é um indicador do poder de compra dos indivíduos e por isso tem larga aplicação nas pesquisas de mercado.

Não há dúvidas que esse tipo de procedimento consegue diferenciar as famílias representadas nesses estudos, mas elas são pouco estáveis e difíceis de serem comparadas temporalmente e espacialmente. Essas evidências de riqueza são muito sensíveis às diferenças regionais, às mudanças nos padrões de consumo das famílias, além de não discriminarem o valor do bem consumido. Itens como televisão a cores e telefone celular estão muito disseminados junto às famílias urbanas e a capacidade de distinção que possuíam há alguns anos já não é tão forte. Novos itens, inicialmente raros e caros, rapidamente se popularizam, como os DVDs e máquinas de lavar roupas. A simples posse de automóvel, cada vez mais comum nos centros urbanos, não permite discriminar as grandes diferenças no valor desse bem. O mesmo vale para modelos de televisores em cores, máquinas de lavar etc. Registram-se também problemas na qualidade dessa informação, conforme discutido por Soares e Andrade (2006), que mostram inconsistências nas respostas sobre bens de conforto doméstico fornecidas pelo mesmo grupo de estudantes (das mesmas escolas) em dois surveys diferentes.

Uma alternativa - não muito comum nas pesquisas sociais, mas que merece ser analisada - é a estimação da renda indiretamente de acordo com o local de residência da família. Essa metodologia é mais usual nas pesquisas em saúde pública (KRIEGER et al, 1997). O pressuposto é que as moradias compartilham uma vizinhança que pode ser caracterizada em termos de sua composição social e econômica. O estudo das áreas vizinhas permite demarcar populações relativamente homogêneas com respeito a características que estruturam diferenças sociais muito importantes.

Em centros urbanos, uma vizinhança pode ser demarcada por setor censitário ou pelo Código de Endereçamento Postal, o CEP. Os setores censitários são áreas geográficas definidas pelo IBGE, obedecendo a critérios de operacionalização da coleta de dados, de tal maneira que abranjam uma área que

possa ser percorrida por um único recenseador em um mês, e que possuam em torno de 250 a 350 domicílios (em áreas urbanas). O setor censitário é a menor unidade para o qual o IBGE fornece informações socioeconômicas, tais como renda média dos chefes de família, porcentagem de chefes de família com nível superior, número de domicílios por tipo etc⁷. É de se presumir, portanto, que a mensuração de conceitos como proximidade ou vizinhança seja mais precisa em nível de setor censitário.

No entanto, localizar nos setores censitários os endereços das famílias representadas em um survey, embora seja possível, não é uma tarefa das mais simples, porque os indivíduos, naturalmente, não sabem informar qual é o número do setor censitário onde está localizado o seu domicílio. Esta tarefa complexa cabe ao pesquisador interessado em utilizar essa informação. O CEP do endereço, uma informação mais facilmente obtida por meio de questionários, constitui uma alternativa para classificar a composição social e econômica de áreas vizinhas. Mas, antes de utilizar essa informação com dados empíricos, é necessário ligar cada número de CEP com os números dos setores censitários que o CEP faz borda ou atravessa. Essa tarefa, também muito exaustiva, pode ser realizada de forma mais fácil se forem utilizados recursos computacionais de georeferenciamento. Assim, os mapas dos setores censitários podem ser sobrepostos aos mapas (trajetos) dos CEPs e gerar bancos de dados com essas informações compartilhadas. Como frequentemente um CEP está ligado a mais de um setor, deve-se atribuir aos CEPs a média das informações sociais e econômicas dos setores censitários a eles ligados. Essa informação pode então ser operacionalizada como uma variável ordinal por faixas de salários mínimos que sejam significativas para discriminar estratos sociais.

Essa forma de operacionalizar a informação sobre a renda foi incorporada neste trabalho. Embora se possa criticar o uso de uma informação agregada da área de residência para atribuir um rendimento à família ou ao indivíduo, argumenta-se que esta opção é mais estável como evidência de uma faixa de rendimento das famílias do que as medidas indiretas da renda pelo consumo de bens. Além do mais, como já foi dito, para a construção da escala de NSE proposta neste trabalho, esta informação não será utilizada isoladamente, mas em conjunto com outros dados de nível individual (educação e ocupação), a fim de aumentar reciprocamente o seu poder de discriminação.

Em relação à ocupação, se, por um lado, há um consenso entre os sociólogos sobre a centralidade dessa informação para a descrição da estrutura

⁷ As informações estatísticas para o nível do setor censitário são disponibilizadas pelo IBGE para municípios acima de vinte e cinco mil habitantes, assim como os mapas de cada setor. É possível fazer download dessas informações ou então comprá-las em CD-Rom na loja virtual, ambos procedimentos disponíveis no sítio do IBGE.

social, por outro lado, não se verifica o mesmo acordo quanto à forma de coletá-la e de classificá-la. Ao contrário da renda, é mais fácil obter respostas fidedignas sobre a ocupação. Nos surveys, os entrevistados usualmente sabem explicar muito bem a natureza do próprio trabalho ou o de outras pessoas da família. Uma vez os dados coletados, as informações sobre ocupação precisam ser codificadas segundo algum esquema válido que possa revelar as diferenças na estrutura ocupacional. Para auxiliar neste processo, o Ministério do Trabalho publicou a Classificação Brasileira das Ocupações – CBO 2002, atualizando a classificação anterior (CBO 1994). A CBO classifica e nomeia 2423 títulos ocupacionais, que recebem códigos de seis dígitos com um alto nível de detalhamento. Em um nível acima, os títulos são agregados em famílias ocupacionais, com quatro dígitos.

Essa nova classificação tem como referência a *International Standard Classification of Occupations – ISCO-88* (ou *Clasificación Internacional Uniforme de Ocupaciones – CIUO-88*), o que torna as estatísticas ocupacionais que utilizam a CBO comparáveis em níveis internacional e nacional⁸. Assim, uma vez as ocupações classificadas de acordo com a CBO, é possível traduzir os códigos para a classificação internacional (ISCO-88) que, por sua vez, podem ser traduzidos para um índice de nível socioeconômico muito utilizado em estudos comparativos internacionais, o ISEI, citado anteriormente (GANZEBOOM e TREIMAN, 1996)⁹.

No entanto, esse processo de coleta de dados ocupacionais e a classificação tomam muito tempo e tem um custo elevado. Para utilizar plenamente a CBO, os questionários devem incluir várias perguntas que caracterizem com propriedade a ocupação (o nome da ocupação, a descrição do tipo de trabalho, o tipo de empresa, o setor, o tipo de vínculo empregatício), o que nem sempre é possível nos surveys. Mesmo que todas as perguntas necessárias sejam incluídas, ou apenas uma ou duas mais importantes que possibilitem a classificação da ocupação em um nível

⁸ Não está ainda completamente solucionada a comparação entre as diferentes classificações ocupacionais utilizadas no país. O IBGE, a partir de 2002, passou a adotar uma nova classificação de ocupações denominada CBO domiciliar. Esta foi definida como uma adaptação da CBO para pesquisas domiciliares e foi criada devido às dificuldades de sua utilização com precisão em pesquisas domiciliares. A CBO domiciliar se mantém idêntica à CBO no nível mais agregado e reagrupa algumas famílias ocupacionais. No entanto, não há uma tábua de conversão entre essas duas formas de classificação, mesmo no nível mais agregado.

⁹ A CBO disponibiliza em seu sítio (www.mtecbo.gov.br/conversao.asp) uma tábua de conversão de seus códigos para os códigos ISCO88. No entanto, essa conversão eletrônica deve ser utilizada com parcimônia, pois há algumas inconsistências, por exemplo, a ocupação de apicultor (6134-05) é convertida para o código ISCO88 3131, que corresponde à ocupação de fotógrafos e operadores de equipamento de gravação de imagem e som. No documento impresso, é possível obter uma conversão mais confiável. A conversão dos códigos ISCO88 para os índices na escala ISEI é facilitada com a utilização de uma sintaxe em SPSS disponibilizada no sítio de um dos autores da escala (home.fsw.vu.nl/~gazeboom/pisa/). Todas essas conversões foram unificadas em um único banco de dados em SPSS por um dos autores deste trabalho durante estágio de pós-doutorado no Departamento de Sociologia da UFMG.

mais agregado da CBO (a família ocupacional), a codificação dessas informações continua sendo uma tarefa bastante complexa¹⁰.

Uma alternativa a essa dificuldade é apresentar categorias ocupacionais agrupadas e hierarquizadas, por exemplo, os seis estratos ocupacionais propostos por Pastore (1979) e atualizados por Pastore e Valle Silva (2000), citados na seção anterior deste trabalho. Naturalmente, não seria viável incluir todas as ocupações em um questionário para que os entrevistados se localizem em um deles. Isso teria que ser feito de forma mais agregada, exemplificando cada estrato com ocupações mais típicas. O vestibular da UFMG, no seu questionário contextual aplicado aos candidatos, inclui um item desse tipo. Os candidatos escolhem a opção que melhor caracteriza a ocupação do pai e da mãe entre seis estratos descritos com exemplos de ocupações. O trabalho de Soares e Andrade (2006), que também produziu uma medida de NSE, utilizou essa variável e comprovou a eficiência da mesma para a discriminação dos perfis de NSE. Isso sugere que pode ser factível o uso de itens fechados, como o do vestibular, para a produção de índices.

Finalmente, para agregar as informações sobre a escolaridade, ocupação e renda em uma única medida de NSE é necessário o uso de técnicas estatísticas multivariadas. Neste trabalho, empregou-se um modelo da Teoria da Resposta ao Item – TRI (HAMBLETON, 1993). A TRI engloba uma série de modelos cujo objetivo principal é a obtenção de medidas de construtos latentes, baseados em indicadores dicotômicos e/ou ordinais. Especificamente, utilizou-se o modelo de Samejima (1969), apropriado para itens com respostas graduadas, ou ordinais¹¹. Essa ferramenta de análise, mais conhecida no Brasil para a produção de proficiências de alunos que respondem a testes de avaliação educacional, foi utilizada por Soares. (2005) e Soares e Andrade (2006) para produzirem escalas de NSE, embora esses trabalhos tenham utilizado outros indicadores. Uma vantagem adicional da TRI em relação aos métodos convencionais é a possibilidade de estimar a medida do NSE mesmo com dados incompletos para algum dos indicadores, pois esse processo é viável mesmo com apenas alguns dos itens respondidos.

¹⁰ Por ter uma idéia, Hauser e Warren (1997) estimam que para coletar as informações necessárias para codificar a ocupação de um indivíduo segundo a classificação do Censo norte-americano são necessárias no mínimo cinco perguntas, que tomarão cerca de um minuto e meio para o seu registro em uma entrevista. Para o tratamento dessa informação, um codificador experiente conseguirá codificar no máximo 10 entradas em uma hora de trabalho.

¹¹ Para leitores interessados na descrição do modelo Samejima, recomenda-se a leitura de Van der Linden e Hambleton (1996), capítulo 5.

Descrição dos dados e procedimentos metodológicos

Neste trabalho, foram utilizados os dados da Pesquisa GERES do Polo Belo Horizonte obtidos nas três primeiras ondas, totalizando informações de 4636 alunos¹². Essa base de dados oferece oportunidade para a implementação dessa proposta de medida do nível socioeconômico porque possui dados de excelente qualidade para esse objetivo. A forma como as variáveis 'educação', 'ocupação' e 'renda' foram consideradas está descrita a seguir.

O item sobre educação dos pais foi incluído no questionário como uma pergunta fechada, com cinco categorias: (1) Nunca estudou ou não terminou 4ª série; (2) Terminou 4ª série do Ensino Fundamental; (3) Terminou 8ª série do Ensino Fundamental; (4) Terminou o Ensino Médio; e (5) Terminou a Faculdade. Na escala de NSE aqui proposta utilizou-se apenas a escolaridade da mãe, porque o efeito da mãe na escolarização dos filhos é considerado maior que o do pai (BUCHMANN, 2002).

Quatro perguntas abertas foram incluídas no questionário GERES para investigar a ocupação dos pais: (1) Qual é a principal ocupação da mãe (ou responsável) do(a) aluno(a)?; (2) O que a mãe (ou responsável) do aluno(a) faz em sua principal ocupação?; (3) Qual é a principal ocupação do pai (ou responsável) do(a) aluno(a)?; e (4) O que o pai (ou responsável) do aluno(a) faz em sua principal ocupação?¹³. Essas perguntas foram codificadas de acordo com a Classificação Brasileira das Ocupações - CBO2002, em duas variáveis: a primeira, com quatro dígitos, para classificar a ocupação no nível da família; e a segunda com dois dígitos, para uma classificação mais detalhada no nível do título ocupacional. Além disso, foram definidos códigos para as situações do pai e/ou mãe que estão fora do mercado de trabalho, como donas-de-casa, aposentados, estudantes e desempregados.

¹² A razão deste estudo se limitar aos dados GERES pólo Belo Horizonte deve-se a impossibilidade de preparar os dados necessários para todos os pólos. Conforme será descrito a seguir, a variável indicadora de renda familiar foi indiretamente medida por meio de informações externas, obtidas junto ao IBGE. Mas essa metodologia poderá ser replicada pelos outros pólos GERES.

¹³ A formulação completa das perguntas é a seguinte: (1) Qual é a principal ocupação da mãe (ou responsável) do(a) aluno(a)? (Por exemplo: enfermeira, professora, empregada doméstica, vendedora, dentista, secretária, dona de casa). Se não estiver trabalhando no momento, conte-nos sobre a última ocupação principal dela; (2) O que a mãe (responsável) do(a) aluno(a) faz em sua principal ocupação? (Por exemplo: cuida de doentes em um hospital; dá aula; trabalha em casa de família; vendedora de uma loja; atende pacientes no consultório; secretária de uma empresa; cuida do lar). Por favor, descreva abaixo o tipo de trabalho que ela faz nessa ocupação; (3) Qual é a principal ocupação do pai (ou responsável) do(a) aluno(a)? (Por exemplo: professor, pedreiro, motorista, médico, comerciante, militar). Se não estiver trabalhando no momento, conte-nos sobre a última ocupação principal dele; e (4) O que o pai (responsável) do(a) aluno(a) faz em sua principal ocupação? (Por exemplo: dá aulas para alunos de quinta a oitava séries; constrói casas; dirige ônibus; atende pacientes no consultório; dono de uma loja; tenente do exército). Por favor, descreva abaixo o tipo de trabalho que ele faz nessa ocupação.

Para o tratamento dessa informação, já codificada para a CBO, inicialmente optou-se por utilizar o *International Socio-economic Index*, o ISEI, de status ocupacional, desenvolvido por Ganzeboom e Treiman (1996) para pesquisas comparativas internacionais. Mas para utilizar essa escala no modelo estatístico proposto, os valores atribuídos do ISEI deveriam ser transformados em uma variável ordinal com poucas categorias. Isso poderia ser feito com o uso da técnica de análise de *Cluster* para definição de grupamentos homogêneos. No entanto, essa opção não teria uma validade à luz das teorias discutidas nesse trabalho, já que os grupos seriam definidos empiricamente com os dados disponíveis. Por esse motivo, optou-se por hierarquizar as ocupações de acordo com os estratos ocupacionais proposto por Pastore e Valle Silva (2002). Esses autores propuseram seis estratos, sendo que o primeiro compreende ocupações predominantemente rurais não qualificadas e o segundo, trabalhadores urbanos não qualificados. Como a amostra GERES inclui apenas escolas urbanas, neste estudo os estratos 1 e 2 foram agregados. Assim, foram definidos cinco estratos ocupacionais: (1) baixo; (2) médio-inferior; (3) médio-médio; (4) médio-superior; e (5) alto. Para alocar as ocupações dentro desses estratos, considerou-se o sistema de classificação da própria CBO (grupos, sub grupos, famílias e ocupações) e também a análise substantiva dos títulos ocupacionais e dos estratos. As ocupações típicas dentro de cada estrato estão descritas no trabalho anteriormente citado¹⁴. Os pais que se encontram fora do mercado (donas-de-casa, desempregados, aposentados, estudantes) não foram classificados e, para esses casos, foi atribuído o código de dado ausente para esse indicador. Ressaltando, como dito anteriormente, que a inexistência de informação sobre a ocupação sobre qualquer um dos outros indicadores não inviabiliza o cálculo do NSE para esses casos, o que é uma das vantagens da TRI. Isso só ocorreria se não houvesse informação para os três indicadores.

Outra decisão importante em relação à ocupação foi a escolha de uma das ocupações, do pai ou da mãe, dado que ambas foram coletadas na pesquisa GERES. Tradicionalmente, as pesquisas sociais medem o status socioeconômico das famílias com a informação da ocupação do pai, assumindo-se, dessa forma, que a posição social da família é determinada pela ocupação do homem adulto e que a ocupação feminina é irrelevante. Essa visão tradicional passou a ser questionada, desde meados dos anos 1970, pelos movimentos feministas e vários

¹⁴ Como dito anteriormente, esses estratos foram definidos a partir de um trabalho original de Pastore (1979) a fim de possibilitar a comparação entre duas décadas. Num primeiro momento, houve uma preocupação se os estratos definidos para ocupações da década de 1970 teriam validade 30 anos depois e se todas as ocupações GERES estariam registradas nesses estudos. Mas, para este trabalho, os resultados obtidos com a alocação das ocupações GERES nos grupos originais pareceram bem satisfatórios no sentido de permitir uma interpretação da escala obtida e principalmente de possibilitar a validação dos resultados em um referencial teórico importante na área da estratificação social.

pesquisadores que argumentam, justificadamente, que a ocupação feminina é importante tanto para a sua própria posição social quanto para a posição social de sua família (AGUIAR et al, 2007). Além do mais, a crescente presença feminina na força de trabalho bem como no papel de “responsável pela família” faz com que a visão tradicional se sustente menos ainda. Neste trabalho, optou-se por considerar a informação do adulto com status ocupacional mais elevado. Esta abordagem assume que o adulto com mais alto status determina o NSE da família como um todo (ERIKSON, 1984), embora se reconheça que essa opção traga o risco de superestimação do NSE de famílias cujo pai tenha uma ocupação manual e a mãe atue no setor de serviços, geralmente com um status mais elevado¹⁵.

O questionário GERES, como é usual nos surveys educacionais, não inclui uma pergunta direta sobre a renda familiar, mas, alternativamente, contém vários itens sobre consumo de itens de conforto e contratação de empregados domésticos que possibilitam estimar uma renda indireta das famílias. Frente à posição crítica assumida em relação à agregação de itens de conforto discutida anteriormente, optou-se por estimar a renda das famílias indiretamente de acordo com o local de residência do aluno. O GERES coleta o endereço da família do aluno, e essa informação pôde ser compatibilizada por meio de um banco de dados gerado com recursos de georeferenciamento, no qual os setores censitários foram alinhados com os endereços de Belo Horizonte¹⁶. Assim, foi possível atribuir a cada endereço dos alunos GERES residentes no município¹⁷ o seu respectivo setor censitário. O passo seguinte foi trazer as informações sociais e econômicas sobre os setores censitários, obtidas no IBGE, para esta base de dados.

O banco de dados gerado revelou um alto poder de discriminação da renda estimada para os alunos GERES. Porém, considerando que essa aplicação tão detalhada seria difícil de ser replicada em outros trabalhos, optou-se por trabalhar com o CEP e não com o endereço do aluno. Mas como frequentemente um CEP está ligado a mais de um setor censitário, neste trabalho, atribuiu-se para cada um dos CEPs a média das informações sociais e econômicas dos setores censitários geograficamente ligados ao CEP. Especificamente, considerou-se a renda média dos responsáveis pelos domicílios (informações disponibilizadas pelo IBGE) localizados

¹⁵ De fato, na pesquisa GERES, a correlação entre o status ocupacional da mãe e do pai, medido pelo ISEI, é moderada (0,5), mas em nível significativo ($p > 0.01$)

¹⁶ Agradecemos à PRODABEL, empresa da Prefeitura de Belo Horizonte, em especial ao Ângelo Rizzo, que disponibilizou um banco de dados com a ligação de todos os endereços do município e seus respectivos CEPs aos setores censitários do IBGE. Isto foi possível graças a um convênio institucional entre a UFMG e a PRODABEL.

¹⁷ Alguns alunos, apesar de frequentarem escolas de Belo Horizonte, moram em cidades da Região Metropolitana. Isto é mais comum nas escolas localizadas nos limites do município. Há também alguns casos de alunos de escolas privadas mais centrais que residem em condomínios no eixo Sul da cidade, que são áreas que concentram famílias de maior poder aquisitivo. Esses foram tratados como casos omissos.

em um determinado CEP. Essas rendas médias foram ordenadas em cinco faixas de salários mínimos (SM), de acordo com os seguintes estratos: (1) Até 2 SM; (2) Mais de 2 a 3 SM; (3) Mais de 3 a 5 SM; (4) Mais de 5 a 10 SM; (5) Mais de 10 SM.

No apêndice, encontram-se as tabelas de distribuição das frequências de cada uma dessas variáveis.

Resultados

Análise dos itens considerados

A análise da qualidade de cada indicador como componente da medida característica latente - NSE - é feita analisando-se a sua curva característica e a sua curva de informação, que podem ser observadas nos gráficos abaixo. Por exemplo, o Gráfico 1 apresenta a curva da variável “escolaridade da mãe”, que possui cinco categorias ordinais. No eixo das abscissas está colocado o nível socioeconômico dos alunos e no eixo das ordenadas está a probabilidade de uma pessoa com dado NSE estar em cada uma das categorias. Pode-se notar que as alternativas 1 e 5 (respectivamente, a mãe que nunca estudou ou não completou a 4ª série e a mãe que completou o ensino superior) correspondem apenas às famílias dos extremos da escala. Além disso, à medida que aumenta o NSE, aumenta a probabilidade de a mãe ter completado o ensino superior e diminui a probabilidade das opções intermediárias (2, 3 e 4, respectivamente).

O Gráfico 2 apresenta a curva de informação do item “escolaridade da mãe”. Esta é uma medida de precisão e mostra o quanto cada uma das posições na escala de NSE está sendo bem estimada. Nota-se que a precisão máxima está no valor zero, e a contribuição da escolaridade da mãe para a estimativa do NSE continua razoável entre os valores -1 e 2. Nos extremos, a contribuição da informação desse item decresce rapidamente. A mesma leitura deve ser feita com os outros gráficos (3 a 6).

As três curvas características mostram que a divisão das variáveis consideradas em cinco categorias ordinais permitiu discriminar bem perfis de famílias ao longo da escala do NSE. As curvas de informação da escolaridade e renda mostram que há menos precisão das estimativas para as famílias com menor NSE (>-2), mas a variável renda apresenta uma curva de informação com melhor cobertura em todo o espectro do NSE (Gráfico 6).

Gráfico 1

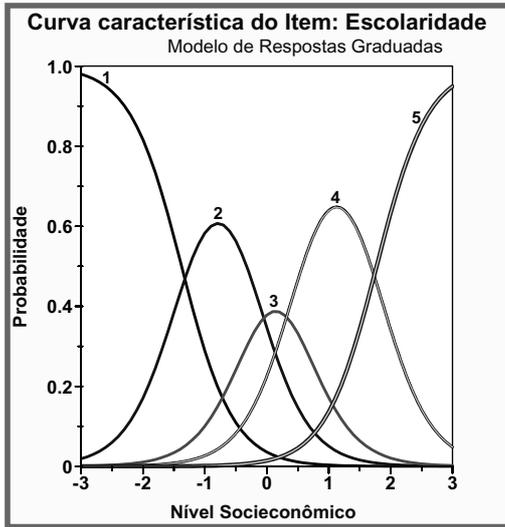


Gráfico 2

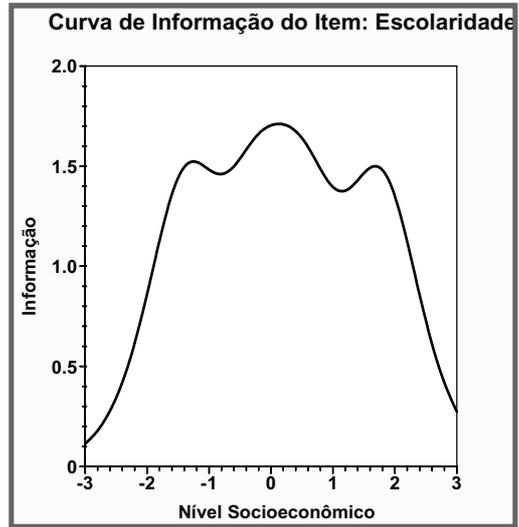


Gráfico 3

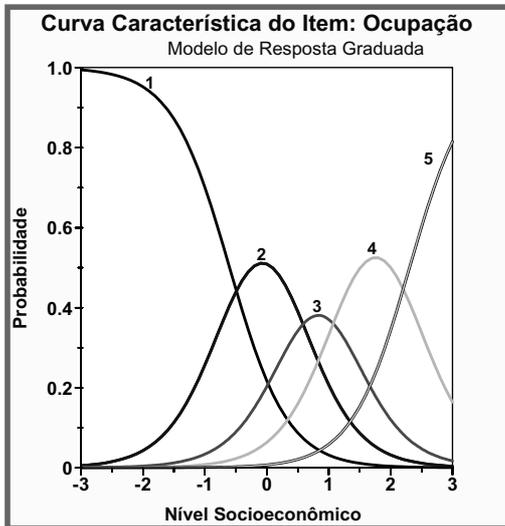


Gráfico 4

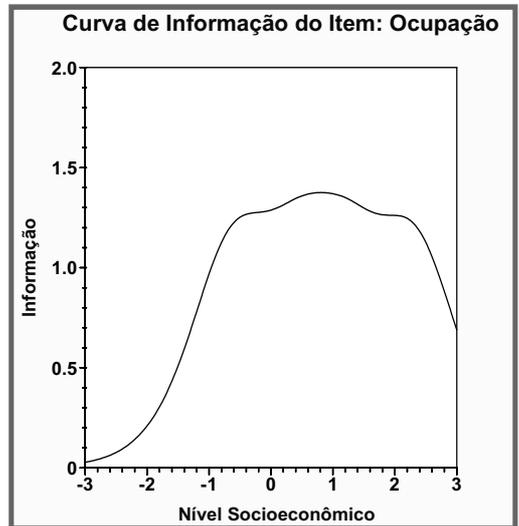


Gráfico 5

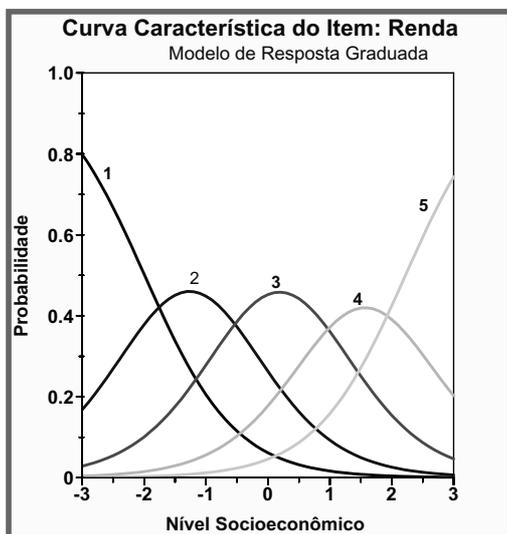
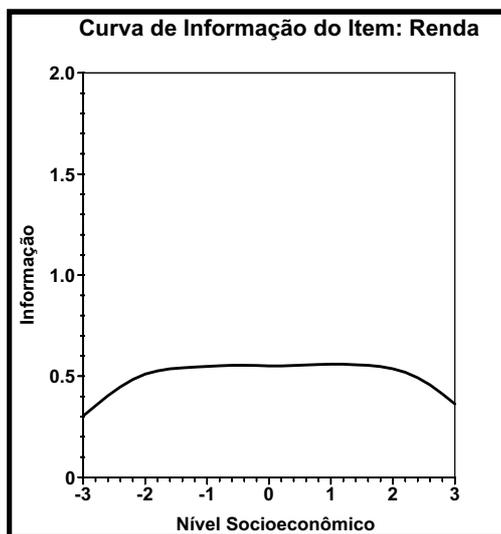


Gráfico 6



A escala de Nível Socioeconômico

A análise dos itens mostrou a adequação da escolaridade, ocupação e renda para a estimação do nível socioeconômico das famílias. Calculou-se então uma medida de NSE para cada um dos alunos da pesquisa considerada. O Gráfico 7 apresenta o histograma dessa medida para 4636 alunos incluídos neste estudo. Há maior concentração de alunos com baixo NSE, o que é esperado já que mais de dois terços das escolas da amostra GERES são públicas, onde a maioria dos alunos vêm de famílias com baixo NSE. De qualquer forma, essa representação gráfica não corresponde à distribuição de alunos desta faixa etária em Belo Horizonte, porque a amostra não é proporcional à população. Esta correção deverá ser feita com a ponderação da amostra quando os dados dessa pesquisa forem analisados em outros estudos, por exemplo, para a análise do valor agregado das escolas.

O Gráfico 8 mostra o NSE dos alunos discriminados por dependência administrativa da escola em que eles estão matriculados. Também, como esperado, observa-se um gradiente entre as redes de ensino, sendo que os alunos das escolas municipais e estaduais têm NSE mais parecidos entre si e nos níveis mais baixos. As diferenças entre as redes são marcantes, porém, é digno de nota que há uma grande variação dentro das redes, tanto com a presença de alguns alunos de alto NSE nas redes públicas quanto de alunos com maior NSE na rede privada.

Gráfico 7

**Histograma do NSE dos alunos
GERES - Polo Belo Horizonte**

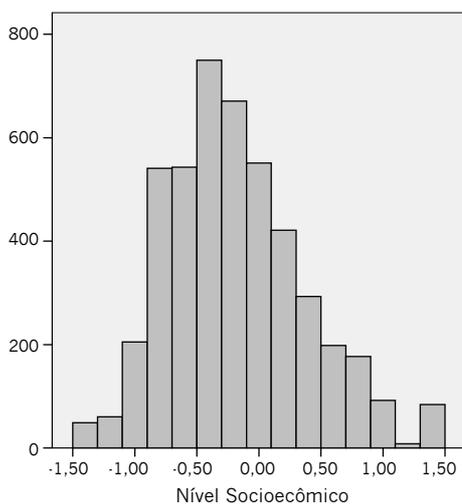
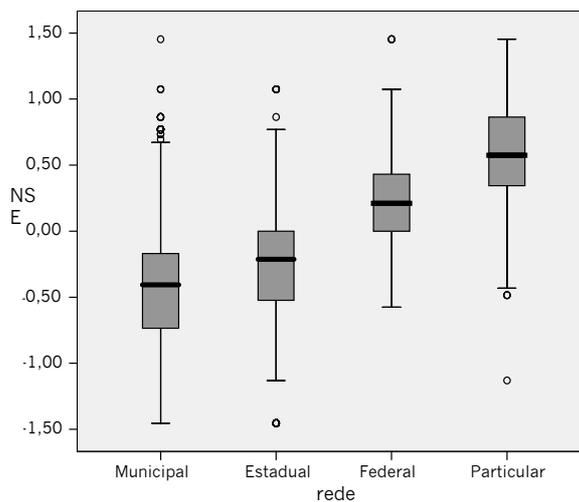


Gráfico 8

**Varição do NSE por
dependência administrativa**



A interpretação da escala do Nível Socioeconômico

A Tabela 1 mostra a distribuição das categorias de cada um dos fatores considerados - escolaridade, ocupação e renda - segundo quintis da escala do NSE produzida com os dados GERES. A caracterização desses grupos se torna bastante significativa quando eles são comparados com a distribuição geral desses fatores.

Tabela 1
Interpretação da Escala de NSE segundo quintis de sua distribuição

	QUINTIS DO NSE					Total
	1	2	3	4	5	
ESCOLARIDADE DA MÃE						
Mãe nunca estudou ou não terminou a 4ª série do EF	47,3%	14,8%	1,8%	1,1%		12,7%
Terminou 4ª série do EF	48,2%	53,0%	47,3%	10,5%	,5%	31,8%
Terminou 8ª série do EF	3,3%	26,2%	38,5%	32,5%	8,1%	21,6%
Terminou o Ensino Médio	1,2%	6,0%	12,4%	54,4%	53,4%	25,5%
Mãe terminou a faculdade				1,4%	38,0%	8,4%
Total	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
ESTRATO OCUPACIONAL						
Baixo	84,6%	56,6%	11,4%	3,3%	,7%	31,2%
Médio-inferior	14,5%	35,8%	67,7%	41,9%	2,9%	32,4%
Médio-médio	,7%	6,9%	18,1%	37,9%	28,2%	18,2%
Médio-superior	,1%	,6%	2,6%	16,6%	46,5%	13,6%
Alto			,2%	,3%	21,7%	4,7%
Total	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
RENDIMENTO						
Até 2 SM	41,3%	10,9%	1,4%	,8%		10,5%
Mais de 2 a 3 SM	48,3%	39,3%	32,4%	13,2%	,8%	26,8%
Mais de 3 a 5 SM	8,5%	41,6%	53,8%	52,0%	18,3%	35,1%
Mais de 5 a 10 SM	2,0%	7,6%	11,0%	29,5%	43,8%	18,8%
Mais de 10 SM		,6%	1,4%	4,5%	37,1%	8,9%
Total	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fonte: produção própria a partir dos dados da pesquisa GERES – Polo Belo Horizonte (3 ondas)

Por exemplo, observa-se na coluna Total que 12,7% das mães desta amostra nunca estudaram ou não terminaram a 4ª série. No primeiro quintil do NSE há quase quatro vezes mais mães nessa categoria (47,3%). Nos outros quintis a frequência de mães nessa categoria é muito pequena, quase nula ou ausente. No outro extremo, a frequência de mães com curso superior não é muito alta na distribuição total da amostra (8,4%). Mas mães com esta característica distinguem o grupo reunido no último quintil (frequência de 38%), já que no quarto quintil apenas 1,4% das mães tem curso superior e não há registro de nenhuma nos outros quintis. A mesma leitura deve ser feita na segunda e terceira parte da tabela, em relação às variáveis ocupação e renda. Os valores mais relevantes de cada categoria das variáveis foram destacados nas tabelas.

Assim, pode-se dizer que na amostra GERES - polo Belo Horizonte, as famílias classificadas com os valores do NSE mais baixos, tipicamente, são compostas por mãe com baixo nível de escolaridade (47,3% nunca estudaram ou não terminaram a 4ª série; 48,2% terminaram a 4ª série). A grande maioria dos responsáveis (pai ou mãe) tem ocupação classificada no estrato mais baixo (84,6%), por exemplo, empregadas domésticas e trabalhadores braçais sem especialização. Essas famílias vivem em áreas cuja renda média dos responsáveis pela família está entre 2 a 3 SM (48,3%), mas um grande percentual (41,3%) recebe no máximo 2 salários mínimos.

O segundo quartil se distingue do anterior principalmente por que há mais mães com escolaridade até a 4ª série (53%) e porque vivem em locais com nível de renda um pouco mais elevado (39,3% vivem em áreas cuja renda média do responsável na faixa de 2 a 3 SM; e 41,6% em áreas com renda na faixa de 3 a 5 SM). Porém, ainda há uma grande concentração de responsáveis em ocupações do estrato mais baixo (56,6%).

No quintil intermediário, há uma frequência alta de mães que terminaram a 4ª série (47,3%), assim como daquelas que terminaram a 8ª série (38,5%). Mas este é o grupo mais típico das ocupações pertencentes ao estrato médio-inferior (67,7%), como motoristas, pedreiros e mecânicos de veículos. As famílias vivem em áreas de renda entre 3 a 5 SM (53,8%), mas também em bairros mais pobres, com renda entre 2 a 3 SM (32,4%).

No quarto quintil, encontra-se uma alta frequência de mães que terminaram o ensino médio (54,4%), mas ainda é alto o percentual das que só terminaram o ensino fundamental (32,5%). Os responsáveis se ocupam de atividades muito variadas, classificadas principalmente nos estratos médio-inferior (41,9%) e médio-médio (37,9%) - caracterizado, por exemplo, por auxiliares administrativos e de escritórios, reparadores de equipamentos, praças das forças armadas e comerciários. As famílias vivem em áreas com renda entre 3 a 5 SM (52%) e entre 5 a 10 SM (16,6%).

No último quintil, encontram-se principalmente as famílias cujas mães têm curso superior (38%), os responsáveis têm ocupação no estrato médio-superior (46,5%), que reúne ocupações tais como diretores, assessores e chefes no serviço público, administradores e gerentes, médio-médio (28,2%) e superior (21,7%), grupo no qual estão os empresários, profissionais de nível superior, oficiais das forças armadas, juízes. Este último estrato somente aparece de forma significativa nesse grupo. Essas famílias vivem nas áreas mais ricas da cidade, nos bairros cujo responsável tem renda entre 5 a 10 SM (43,8%) ou mais de 10 SM (37,1%).

Conclusão

Neste trabalho, apresentou-se uma proposta para medir o nível socioeconômico das famílias a partir de dados empíricos obtidos em surveys, tomando como referência os dados da Pesquisa GERES - Polo Belo Horizonte.

Do ponto de vista operacional, o nível socioeconômico - um construto teórico não diretamente observado - foi medido a partir de três indicadores: a escolaridade, a ocupação e a renda. Esses foram definidos à luz da literatura sociológica com o objetivo de produzir uma medida que fosse teoricamente fundamentada. Para a agregação dos fatores (escolaridade, ocupação e renda) foi utilizado um modelo da Teoria da Resposta ao Item - TRI. Este método oferece vantagens em relação às abordagens tradicionais de construção de escalas (como a análise fatorial) porque permite trabalhar com indicadores categóricos (dicotômicos ou ordinais) e tem uma forma natural para lidar com os dados ausentes. O resultado obtido foi uma escala hierárquica que mede o NSE a partir de uma combinação ponderada das diferentes combinações dos fatores escolaridade, ocupação e renda.

A revisão da literatura mostrou a centralidade da variável ocupação na definição das posições sociais. No entanto, na aplicação empírica das teorias, verifica-se que é muito difícil derivar uma medida de status socioeconômico sem informações sobre a educação e o rendimento associados a cada posto ocupacional, além de frequentemente serem também necessárias outras informações como o sexo e a idade dos ocupantes desses postos (GANZEBOOM e TREIMAN, 1996; PASTORE e VALLE SILVA, 2000).

Na aplicação aos dados GERES, a medida de NSE apresentada neste trabalho foi baseada nas informações sobre a escolaridade da mãe do aluno, o estrato ocupacional de maior status entre pai e mãe e a renda, e a renda do responsável pela família, considerando como tal a renda média dos responsáveis pelas famílias que vivem na mesma área onde o aluno reside.

A novidade deste trabalho foi quanto à forma de operacionalização dessas variáveis, especialmente a ocupação e a renda. Mas os resultados mostraram que é

possível utilizar essas variáveis em pesquisas do tipo survey (mesmo com questionários autoaplicados).

A ocupação, coletada na pesquisa GERES com perguntas abertas, foi recodificada segundo estratos ocupacionais propostos por Pastore e Silva (2000). Com base nesse exercício, recomenda-se que a ocupação dos entrevistados ou dos responsáveis (no caso de alunos) sempre faça parte dos questionários das pesquisas sociais. Isso traria uma informação importante para a produção de uma medida de NSE. Reconhece-se, no entanto, que perguntas abertas significam custos mais elevados para as pesquisas. Mas, se a finalidade da pergunta for a elaboração de um índice, a alternativa do Vestibular da UFMG (um item fechado pré-codificado segundo estratos ocupacionais exemplificados por ocupações típicas) pode ser considerada adequada, conforme discutido por Soares e Andrade (2006). Porém, não está claro como isso poderia ser generalizado para qualquer tipo de público, por exemplo, nos questionários respondidos por alunos mais novos, que certamente teriam dificuldade de alocar a profissão de seus pais dentro de um pequeno grupo de possibilidades. Nesse caso, a melhor alternativa é a aplicação dos questionários diretamente aos pais, como é feito na Pesquisa GERES¹⁸.

A opção por medir a renda indireta através do local de moradia do aluno mostrou-se bem interessante para este estudo. Mas se reconhece a dificuldade de generalização dessa metodologia para todos os tipos de pesquisa. As bases de informações por setores censitários estão disponíveis para municípios acima de vinte e cinco mil habitantes. Além disso, a utilização dessa informação depende de um grande esforço de ligação dos endereços aos setores censitários. A alternativa mais viável é a ligação dos setores aos Códigos de Endereçamento Postal (o CEP), utilizando-se preferencialmente para isso uma base de informações georeferenciadas dos setores e dos CEPs¹⁹.

Neste trabalho, experimentou-se as duas formas, i.e., a ligação dos setores censitários diretamente aos endereços e a ligação de atributos médios dos setores censitários que fazem limites ou atravessam os CEPs do município. A primeira resultou em uma estimativa mais fina da renda média do local onde o aluno vive, porém, exigiu muito tempo de trabalho e é uma metodologia de difícil replicação, pois essa localização exata somente foi possível devido à disponibilidade dos dados por uma fonte externa. A segunda resultou em perda de informação, porque,

¹⁸ Isso também pode ser feito em pesquisas em larga escala. Por exemplo, na Prova São Paulo, avaliação realizada no município de São Paulo junto aos alunos matriculados na rede na 2ª, 4ª, 6ª e 8ª séries, os pais dos alunos respondem aos questionários contextuais. Nessa situação, é possível obter informações mais fidedignas sobre a ocupação dos pais.

¹⁹ É também possível fazer a ligação dos CEPs aos setores a partir da consulta aos mapas dos setores e uma base de dados com todas as ruas e respectivos CEPs de um município. Dessa forma, gera-se um banco de dados no qual todos os CEPs estarão ligados a um ou mais setores censitários e vice-versa. Mas, certamente, o recurso de georeferenciamento facilita sobremaneira esse trabalho.

diferentemente dos setores censitários, o CEP cobre áreas bem maiores, menos homogêneas, uma vez que a lógica do endereçamento postal (facilitar o trabalho de distribuição de correspondências) não é a mesma do setor (definir áreas para o trabalho de recenseamento). Mas um teste realizado com os dois resultados mostrou que a correlação da renda atribuída para a composição da vizinhança por meio do setor censitário (mais preciso) ou pelo CEP é muito elevada (0,87), o que justificou a opção pela segunda, mais viável.

Uma dificuldade adicional do uso dessa metodologia é que muitas cidades ainda não possuem uma estrutura de CEP por rua. Portanto, embora se reconheça que essa informação seja de melhor qualidade para inferir sobre a renda indireta da família do que as alternativas usuais (por exemplo, a posse de bens de conforto que medem poder de compra), a mesma metodologia poderá ser utilizada apenas em pesquisas realizadas em grandes centros urbanos. A generalização dessa metodologia depende, sobretudo, do sucesso em se obter a informação do endereço dos entrevistados, mesmo que incompleto (por exemplo, apenas a rua ou apenas o CEP). Mas também não está claro se isso seria aceito em qualquer contexto, por exemplo, nas pesquisas para avaliação de desempenho escolar.

Uma última questão em relação à informação sobre a renda refere-se às categorias definidas. A variável sobre renda dos responsáveis pelas famílias foi categorizada em cinco faixas, sendo que a primeira compreende as famílias cujos responsáveis têm uma renda de até dois salários mínimos e a última faixa, com renda acima de dez salários mínimos. Essas faixas foram definidas com referências às pesquisas de opinião pública de uma forma geral, mas também para evitar categorias com um número de casos muito reduzidos ou cuja curva (na análise dos itens) fosse sobreposta por outra. No entanto, se reconhece que essas faixas extremas talvez pudessem ser divididas para melhor captar os diferentes graus de pobreza das famílias na faixa mais baixa e de riqueza na faixa mais alta.

Apesar dessas ressalvas, o resultado obtido mostrou que o nível socioeconômico calculado a partir das informações sobre a escolaridade, ocupação e renda discrimina muito bem as famílias GERES ao longo da escala de NSE. Esta, no entanto, não é muito diferente da escala anterior calculada para os alunos da pesquisa GERES²⁰. De fato, observa-se uma validade convergente entre as duas escalas de NSE, no sentido que medem o construto nível socioeconômico com alta correlação entre si²¹. Assim, em qualquer estudo com os dados GERES para analisar questões substantivas relacionadas ao valor agregado das escolas ou

²⁰ Nota técnica sobre o cálculo do nível socioeconômico para os alunos GERES disponível no sítio <www.geres.ufmg.br> (acesso em nov. 2007).

²¹ A correlação entre as escalas de NSE obtidas pelos dois métodos é de 0,8.

outros temas, espera-se que o efeito do NSE se manifeste da mesma forma utilizando-se uma ou outra escala. Contudo, procurou-se, com este estudo, garantir uma maior validade externa para a escala de NSE, no sentido de que ela possa ser generalizada a outros contextos de pesquisa e com maior consistência teórica.

Uma possível crítica à medida do NSE proposta é a impossibilidade de interpretação nos modelos de análise do efeito único das dimensões agregadas no índice. Educação, renda e ocupação constituem categorias analíticas muito sólidas nas pesquisas sociais, mas elas perdem seu status teórico quando incorporadas em uma medida única. Essa crítica pode ser mais ou menos importante, dependendo do objetivo do estudo. Nas pesquisas sobre o efeito das escolas (ALVES e FRANCO, 2008), por exemplo, é imprescindível uma medida de controle sobre os fatores externos às escolas (o nível socioeconômico dos alunos, principalmente), pois somente assim é possível isolar o impacto das escolas para o resultado dos alunos. Portanto, o foco desse tipo de estudo é, em última análise, a escola e não o efeito isolado de variáveis externas, que entram no modelo como controle. Nesse caso, a medida de NSE proposta tem a vantagem de contornar o problema de ausência de resposta que, comumente, ocorre na coleta de informações das variáveis isoladas. Como dito anteriormente, a TRI pode estimar a medida do NSE mesmo com dados incompletos para alguns dos indicadores. Dessa forma, evita-se a perda de casos na análise da questão substantiva em foco.

Outra crítica é o fato de que, no índice construído, as variáveis educação e renda impactam duplamente o índice, pois são consideradas tanto como indicadores isolados, como também para a criação da hierarquia das ocupações, embora, neste caso, a partir de outra fonte de dados. Isso gera associação entre os indicadores de ocupação e os de educação e renda. No entanto, a técnica usada para agregação dos indicadores considera naturalmente as associações entre os indicadores e mantém na medida final apenas a contribuição específica de cada um.

Finalmente, reconhecendo que o NSE constitui um determinante-chave para a análise de dados em todas as pesquisas sociais, espera-se que este trabalho possa fornecer ferramentas metodológicas para as pesquisas empíricas nesse campo. Naturalmente, a proposta de medida do nível socioeconômico apresentada neste artigo poderá ser aperfeiçoada com novos estudos, por exemplo, o desenvolvimento de escalas mais atuais para a classificação das ocupações e a utilização das bases de dados do IBGE por um maior número de pesquisadores.

Referências Bibliográficas

AGUIAR, N; D. FERNANDES e J. A. NEVES. Mobilidade Social Feminina. In: AGUIAR, N. (Ed.). *Desigualdades Sociais, redes de sociabilidade e participação política*. Belo Horizonte: Editora UFMG, p.165-180, 2007.

ALVES, M. T. G. e FRANCO, C. A pesquisa em eficácia escolar no Brasil: evidências sobre o efeito das escolas e fatores associados à eficácia escolar. In: BROOKE, N e SOARES, J. F. (Org.). *Pesquisa em eficácia escolar: origem e trajetórias*. Belo Horizonte: Editora UFMG, p. 482-500, 2008.

BLAU, P.M. e DUNCAN, O.D. *The American Occupational Structure*. New York: Wiley. 1967.

BOLLEN, K. A.; GLANVILLE, J. e STECKLOV, G. "Socioeconomic Status and Class in Studies of Fertility and Health in Developing Countries". *Annual Review of Sociology*, v.27, p.153-185, 2001.

BUCHMANN, C. Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges. In: PORTER, A. e GAMORAN, A. (Ed.). *Methodological Advances in Cross-National Surveys of Educational Achievement*. Washington, DC: National Academy Press, p.150-197, 2002.

CIRINO, P. T.; C. E. CHIN; R. A. SEVCIK; M. WOLF; M. LOVETT e R. D. MORRIS. Measuring Socioeconomic Status: Reliability and Preliminary Validity for Different Approaches. *Assessment*, v.9, n.2, June, p.145-155, 2002.

DUNCAN, O. D. A Socioeconomic Index for All Occupations. In: REISS, A.J. (ed). *Occupations and Social Status*. New York: Free Press. p. 109-138, 1963.

ENSMINGER, M. E.; FORREST, C. B.; RILEY, A. W.; KANG, M.; GREEN, B. F.; STARFIELD, B. e RYAN, S. A. "The Validity of Measures of Socioeconomic Status of Adolescents". *Journal of Adolescent Research*, v.15, n.3, p.392-419, 2000.

ENTWISLE, D. R. e ASTONE, N. M. "Some Practical Guidelines for Mearuring Youth's Race/Ehnicity and Socioeconomic Status". *Child Development*, v.65, p.1521-1540, 1994.

ERIKSON, R. "Social class of men, women and families". *Sociology*, n.18, p. 500-514, 1984.

ERIKSON, R. e GOLDTHORPE, J. "Individual ou Family? Results from Two Approaches to Class Assignment". *Acta Sociologica*, v.35, p.95-105, 1992a.

_____. "The CASMIN Project and the American Dream". *European Sociological Review*, v. 8, n. 3, Special Edition on Social Stratification and Social Mobility, 283-305, 1992b.

FERRÃO-BARBOSA, M. E.; FERNANDES, C. A escola brasileira faz diferença? uma investigação dos efeitos da escola na proficiência em matemática dos alunos da 4a série. In: FRANCO, C. (Ed.). *Promoção, ciclos e avaliação educacional*. Porto Alegre: ArtMed, p.155-172, 2001.

GANZEBOOM, H. B. G.; De GRAAF, P. M. e TREIMAN, D. J. "A Standard International Socio-Economic Index of Occupational Status". *Social Science Research*, v.21, p.1-56, 1992.

GANZEBOOM, H. B.; TREIMAN, D. J. "Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations". *Social Science Research*, n.25, p.201-239, 1996.

GOLDTHORPE, J. H. *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon Press, 1987.

HAMBLETON, R. K. Principles and selected applications of Item Response Theory. In: LINN, R. L. (Ed.). *Educational measurement*. 3. ed. Washington, DC: American Council on Education and the National Council on Measurement in Education, p.147-200, 1993.

HAUSER, R. M.; WARREN, J. R. Socioeconomic Index of Occupational Status: A Review, Update, and Critique. In: RAFTERY, A. (ed.). *Sociological Methodology*, Cambridge: Blackwell, 1997, p. 177-298.

KRIEGER, N.; WILLIAMS, D. R. e MOSS, N. E. "Measuring Social Class in US Public Health Research: Concepts, Methodologies, and Guidelines". *Annual Review Public Health*, n. 18, p. 341-378, 1997.

LYNCH, J. W. e KAPLAN, G. A. Socioeconomic Position. In: BERKMAN, L. F. e KAWACHI, I. (Eds.), *Social Epidemiology*, New York: Oxford, p. 13-35, 2000.

MAY, H. A. "Multilevel Bayesian Item Response Theory Method for Scaling Socioeconomic Status in International Studies of Education". *Journal of Educational and Behavioral Statistics*, v.31, n.1, Spring, p.63-79, 2006.

MUELLER, C. W. e PARCEL, T. "Measures of Socioeconomic Status: Alternatives and Recommendations". *Child Development*, v.52, p.13-30, 1981.

NAKAO, K. e J. TREAS. "Updating Occupational Prestige and Socioeconomic Scores: How the New Measures Measure Up". *Sociological Methodology*, v.24, p.1-72. 1994.

OAKES, J. M. e ROSSI, P. H. "The Measurement of SES in Health Research: Current Practice and Steps toward a New Approach". *Social Science & Medicine*, v. 56, p. 769-784, 2003.

OSBORN, A. F. "Assessing the socio-economic status of families". *Sociology*, v.21, n.3, p.429-448, 1987.

PASTORE, J. *Desigualdade e mobilidade social no Brasil*. São Paulo: Editora da Universidade de São Paulo, 1979.

PASTORE, J. e SILVA, N. V. *Mobilidade Social no Brasil*. São Paulo: Markron. 2000

SAMEJIMA, F. Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, New York, v. 17, 1969. Monograph Supplement , n. 4, Part 2.

SANTOS, J. A. F. *Estrutura de Posições de Classe no Brasil: Mapeando, Mudanças e Efeito na Renda*. Belo Horizonte/Rio de Janeiro: Ed. UFMG/IUPERJ, 2002.

_____. "Uma classificação socioeconômica para o Brasil". *Revista Brasileira de Ciências Sociais*, v.20, n.58, p.27-45, 2005.

SCALON, M. C. "Mapeando Estratos: Critérios para Escolha de uma Classificação". *Dados - Revista de Ciências Sociais*, v.41, n.2. 1998.

SIRIN, S. "Socioeconomic Status and academic Achievement: A Meta-Analytic Review of Research". *Review of Educational Research*, v.75, n.3, Fall, p.417-453, 2005.

SOARES, J. F. O efeito da escola no desempenho cognitivo de seus alunos. In: Souza, A. M. (Ed.). *Dimensões da avaliação educacional*. Petrópolis: Editora Vozes, p.174-204, 2005.

SOARES, J. F. e ANDRADE, R. J. Nível socioeconômico, qualidade e equidade das escolas de Belo Horizonte. *Ensaio: Avaliação e Políticas Públicas em Educação*, v.14, n.50, p.107-126. 2006.

SOARES, T. M. "Utilização da teoria da resposta ao item na produção de indicadores sócio-econômicos". *Pesquisa Operacional*, Rio de Janeiro, v. 25, n. 1, 2005 .

TREIMAN, D. J. *Occupational prestige in comparative perspective*. New York: Academic Press, 1977.

VAN DER LINDEN, W. e HAMBLETON, R. K. (ed.). *Handbook of Modern Item Response Theory*. New York: Springer Verlag, 1996.

WHITE, K. R. "The Relation between Socioeconomic Status and Academic Achievement". *Psychological Bulletin*, v.91, n.3, p.461-481, 1992.

WRIGHT, E. O. *Classes*. London. Verso, 1985.

Apêndice
Escolaridade da mãe

	Frequência (N)	Percentual (%)	Percentual sem casos omisos (%)	Percentual acumulado (%)
1.Nunca estudou ou não terminou a 4ª série do EF	553	11,9	12,7	12,7
2.Terminou a 4ª série do EF	1382	29,8	31,8	44,6
3.Terminou a 8ª série do EF	937	20,2	21,6	66,2
4.Terminou o Ensino Médio	1105	23,8	25,5	91,6
5.Terminou a Faculdade	364	7,9	8,4	100,0
Total	4341	93,6	100,0	
Sem informação	295	6,4		
Total	4636	100,0		

Fonte: produção própria a partir dos dados da pesquisa GERES – Pólo Belo Horizonte (3 ondas)

Maria Teresa Gonzaga Alves - mteresagalves@gmail.com
José Francisco Soares - francisco-soares@ufmg.br

Recebido para publicação em setembro de 2008.
Aprovado para publicação em novembro de 2008.