

# UM MODELO PARA ANÁLISE DA ESTRATIFICAÇÃO EDUCACIONAL NO BRASIL

*Nelson do Valle Silva*

Do Laboratório Nacional de Computação Científica, CNPq

*Alberto de Mello e Souza*

Da Faculdade de Educação da UFRJ

Agradecemos os comentários de Antonio C. C. Campino, bem como a ajuda financeira recebida do INEP. Este trabalho contou com a assistência computacional de Déborah Roditi. Falhas e omissões que tenham permanecido são de responsabilidade exclusiva dos autores. Versão preliminar deste trabalho foi apresentada no VI Encontro Brasileiro de Econometria, São Paulo, dezembro 1984.

---

## RESUMO

O artigo considera o efeito da origem social sobre a realização escolar no Brasil, utilizando os dados da PNAD-1976. Existem duas vantagens sobre os demais trabalhos existentes. Primeiro, a especificação logística do modelo é capaz de isolar os efeitos das variáveis independentes das modificações ocorridas nas distribuições marginais das variáveis, o que permite uma comparação dos coeficientes estimados para as transições escolares consideradas. Esta separação não é possível nos demais modelos. A segunda vantagem reside no escopo nacional da amostra, que retrata uma situação recente. Porém, a PNAD-1976 não inclui informações sobre as chamadas variáveis intervenientes ou sobre muitos fatores que atua, do lado da oferta de escolaridade. Apesar disso, foi possível mostrar o efeito declinante das variáveis de origem social sobre as chances de progressão escolar, independentemente de qualquer eventual homogeneização que possa ocorrer com essas variáveis.

## SUMMARY

The aim of this article is to take into consideration the effect of social origin on the achieved level of schooling in Brasil. This study uses data from the Household survey year 1976 (PNAD - 76) and presents two advantages over the preceding studies in the area. First, the logistic specifications of the model permit to isolate the effect of independent variables from modifications occurred in the marginal distribution of the variables. This possibility allows a comparison among the coefficients estimated for the considered levels of schooling. The second advantage bears upon the national scope of the sample. However, PNAD-76 has no information about the so-called intervenient variables, in this sense, it does not refer to many factors at work beside schooling opportunities. Nevertheless, it was possible to show the decreasing effect of social origin variables over the chances of school promotion, regardless of eventual standardization of these variables homogenization.

## INTRODUÇÃO

Na sociedade brasileira, como de resto em todas as sociedades modernas, as escolas desempenham um papel fundamental na seleção e classificação social dos indivíduos. Cada vez mais, é através delas que as famílias conseguem legar a seus filhos as posições que ocupam na hierarquia social. Mas é também através delas que traços sociais básicos, tais como valores e comportamentos, são transmitidos de geração em geração. A escola tem assim simultaneamente duas grandes funções, freqüentemente contraditórias, de selecionar e socializar os jovens membros de uma sociedade, sendo tanto vista como a *grande equalizadora* como a *grande peneira* (Husén, 1975). A estratificação educacional diz respeito à dependência do desempenho do sistema escolar enquanto responsável por aquelas duas funções, em relação à origem social dos alunos que por ele passam. Em outras palavras, diz respeito à relação entre as características de origem sócio-econômica dos alunos na entrada do sistema escolar e as características individuais observáveis na sua saída, bem como aos mecanismos através dos quais essa relação é estabelecida.

Este trabalho pretende ampliar os conhecimentos sobre a estratificação educacional brasileira de duas maneiras distintas. Primeiro utilizando dados recentes, representativos da população brasileira e que cobrem todos os níveis educacionais. Segundo, mediante o emprego de um modelo logístico, cujas propriedades o tornam mais atraente que os modelos de estudos anteriores. Sua organização é a seguinte: na segunda seção são sumariados alguns dos principais estudos feitos, em especial para a realidade brasileira. Em seguida, procede-se a uma discussão sobre as vantagens do modelo logístico. Após, são relacionadas as variáveis de origem social utilizadas e considerados os efeitos da omissão de variáveis intervenientes e de variáveis ligadas à oferta escolar. Na quinta seção, são apresentados os principais resultados das estimativas obtidas. Finalmente, conclui-se com um sumário dos aspectos mais relevantes do trabalho.

## FATORES DETERMINANTES DA ESTRATIFICAÇÃO EDUCACIONAL

Na discussão sobre estratificação educacional existe uma acentuada predominância de fatores causais ligados à demanda de escolaridade. Para os países mais desenvolvidos esta predominância pode ser explicada pela compulsoriedade efetiva atingida pelo ensino de 1º grau, pela comparativamente reduzida diferenciação das escolas públicas e pela diminuta importância da restrição de vagas no acesso aos demais níveis de ensino. Adicionalmente, boa parte dos estudos sobre a estratificação educacional foi feita por sociólogos, cuja ênfase é posta sobre variáveis de origem familiar, que acentuam aspectos do lado da demanda.

Para os países mais pobres, a importância de fatores atuando do lado da oferta não pode deixar de ser ressaltada. Primeiro, porque a inexistência da escola em certos lugares, especialmente na zona rural e na periferia urbana, é bastante comum. Segundo, porque fatores intra-escolares atuam no sentido de dife-

renciar a qualidade do ensino. Esses fatores podem advir da existência de três turnos diários nas escolas, reduzindo a permanência do aluno a menos de três horas por dia ou do congestionamento das salas de aula. Também podem se manifestar através da insuficiência e inadequação do material escolar oferecido pela escola e até mesmo pela atitude dos professores diante de alunos com valores culturais diferentes dos imperantes na escola. A dificuldade de captar diretamente a influência da oferta acaba por fazer com que as variáveis relacionadas à origem familiar também estejam expressando elementos covariantes do lado da oferta.

Iniciaremos a discussão apresentando resultados obtidos nos países mais desenvolvidos. Os primeiros modelos sociológicos consideravam os efeitos de características familiares, como a ocupação ou educação dos pais sobre a escolaridade dos filhos (Duncan, 1967). Posteriormente, como no modelo de Wisconsin, foram introduzidas variáveis intervenientes entre a origem familiar e a realização educacional. Aquele modelo fez um acompanhamento longitudinal dos graduados do 2º grau no estado de Wisconsin (Sewell e Hauser, 1975). Foram as seguintes as variáveis intervenientes utilizadas: capacidade cognitiva; desempenho acadêmico; aspirações ocupacionais dos estudantes e influência dos chamados "outros significantes", como pais, mestres e amigos.

Os resultados obtidos para os países mais desenvolvidos revelam um consenso de que uma parte nada negligenciável, embora modesta, da variância na realização educacional pode ser atribuída à origem social dos indivíduos. Interpretando a evidência para a sociedade americana, Duncan observa que a "relação entre escolaridade e *background* social é fluida o suficiente para que a escolaridade de uma criança não seja estritamente determinada ou mesmo seriamente limitada pelas circunstâncias da família na qual ela nasceu" (Duncan, 1967, p. 371). Esse e outros estudos, como o de Hauser e Featherman (1976) mostram uma marcante estabilidade temporal dos efeitos da origem social sobre a realização educacional dos indivíduos, apesar das profundas mudanças no perfil sócio-econômico da sociedade americana, acompanhadas do aumento da média de escolarização e da redução de sua variância.

A observação geral que emerge é portanto a de que a escola é não só um veículo para a transmissão intergeracional da posição social da família mas também uma via relativamente autônoma para a mobilidade social. Observa-se também que o impacto da origem social do aluno parece diminuir conforme aumenta a escolaridade do mesmo, parcialmente refletindo um processo de seleção que torna o alunado gradativamente mais homogêneo nos níveis educacionais mais elevados.

Porém, como mostram os resultados do modelo de Wisconsin, a maior parte do efeito das variáveis de origem familiar (no caso, ocupação do pai, educação do pai e da mãe, e renda familiar) sobre a realização escolar do filho se dá através das variáveis intervenientes, numa extensão de até 72% daquele efeito. O efeito dessas variáveis é substancial, sobretudo quando comparado com o efeito direto, de cerca de 12% das variáveis de origem social. Ao todo, o modelo

acima é capaz de explicar 88% da covariância entre *status* familiar e realização educacional.

Uma outra regularidade que tem sido constatada é que o processo de estratificação educacional parece ser basicamente o mesmo para as sociedades industrializadas ocidentais, mesmo que possuam sistemas educacionais tão diferentes, como os Estados Unidos e a Inglaterra. Com efeito, Turner (1960) considera os sistemas educacionais desses países como protótipos, sendo os Estados Unidos o exemplo mais nítido de um sistema do tipo competitivo, em que as escolhas educacionais são postergadas o mais possível. Já o sistema inglês é do tipo cooptativo, por exigir, muito cedo, decisões educacionais devido ao sistema rígido de especialização curricular. Porém, Kerckhoff (1974) mostra que apesar dessas diferenças nos sistemas educacionais dos dois países, suas resultantes, quando mensuradas em termos de dependência da realização educacional em relação à origem social do estudante, são bastante parecidas.

Comparando a estratificação educacional nos Estados Unidos e na França, outro exemplo marcante de sistema cooptativo, Garnier e Hout (1975) observam que o processo de estratificação entre os dois países é muito semelhante. A maior desigualdade de oportunidades educacionais na França deve resultar da influência que o professor tem na distribuição dos alunos entre os diferentes currículos, ao contrário da Grã-Bretanha, que utiliza exames universais.

Na formulação dos economistas para a demanda de educação, esta era vista como um bem de consumo e, mais tarde, como um investimento. Embora com um maior poder analítico, nem sempre esta versão da demanda se mostrou superior (Blaug, 1976). Alguns procuraram estudar o atraso escolar e a repetência como decorrência de variáveis como renda e escolaridade dos pais, cor, sexo e local de residência (Conlisk, 1969; e Masters, 1969). Os resultados apontavam a influência crescente dessas variáveis para os maiores níveis de escolaridade.

Outra linha de estudos investigou a relação entre a qualidade da escola e a quantidade da escolaridade, com resultados controversos. Wachtel (1975) não detectou evidência estatística desta relação; porém, Johnson e Starfford (1973) concluem pela sua validade. Gustman e Pidot encontraram uma interação entre qualidade da escola pública, medida pelos gastos correntes por estudante, e a proporção de matrículas na escola pública, resultado de uma transferência de aluno da escola particular para a pública e de uma redução da evasão escolar.

Nos países menos desenvolvidos os estudos estão mais voltados para a explicação do desempenho escolar<sup>1</sup>. A evidência da maioria desses estudos revela a importância predominante do ambiente familiar para os alunos do primário e dos anos iniciais do secundário. Posteriormente, as variáveis relacionadas com a escola ganham ascendência sobre o ambiente familiar, em parte, por este se tornar homogêneo. Como o desempenho escolar define a progressão do aluno, resultados semelhantes devem ser obtidos para a estratificação educacional. Também, dado que a grande seletividade do processo educacional não é aleatória, deve-se esperar que a qualidade da escola esteja associada às características familiares.

No Brasil Lerner (1970), usando dados do Censo Demográfico de 1960, mostrou a importância da transmissão inter-geracional da educação, especialmente na zona rural. Por exemplo, na zona rural a probabilidade de entrar para o primário é 2,5 vezes maior se o chefe da família tiver alguma educação e 3,3 vezes maior se tiver pelo menos o secundário, comparada com o caso de um chefe de família que não foi à escola. Já na zona urbana, neste último caso, a probabilidade é apenas 1,4 vezes maior. Considerando-se a mesma diferença na educação do chefe de família, a probabilidade de concluir o primário na zona rural aumenta de 8,6 vezes e na zona urbana de 1,5 vezes (Lerner, 1970, p. 64). Lerner constatou importantes diferenças na escolaridade de brancos e não-brancos, o que foi observado em dados mais recentes por Silva (1981).

Hansen (1977) fez uso do modelo de Wisconsin em uma amostra obtida no Rio Grande do Sul para explicar as aspirações educacionais e ocupacionais. Constatou a enorme importância da origem social, especialmente quando contrastada com o diminuto efeito da capacidade cognitiva. Porém, pode-se presumir que essa diferença seja atenuada quando se considera a realização educacional.

Um estudo relativo a Belo Horizonte mostra a importância da renda familiar na escolarização dos filhos com mais de 11 anos (Schmidt e Mirand, 1977). A condição de migrante produz um pequeno efeito negativo sobre a escolarização. Mello e Souza (1979) confirmam para o município do Rio de Janeiro a influência crescente da renda familiar para aqueles que já concluíram o primário. Quanto à escolaridade da mãe, sempre teve importância, que vai-se acentuando na medida em que a idade do aluno aumenta. Tamanho da família tem um impacto negativo sobre a escolarização dos filhos, o que significa dizer que famílias maiores reduzem seus gastos *per-capita* em educação. Quando se considerou a origem da renda, verificou-se que os filhos de assalariados recebem uma escolarização mais longa que filhos de autônomos. A constatação semelhante chegou Gouveia (1980), quando observa que, além do nível ocupacional do pai, é importante o tipo de ocupação. Assim, as chances de uma escolarização posterior ao primário são menores para os filhos de comerciantes que para aqueles cujos pais são trabalhadores não-manuais de rotina; da mesma forma, uma maior concentração de indivíduos com o nível superior aparece entre os filhos de profissionais e técnicos de nível superior comparativamente aos filhos de industriais e altos administradores de bancos.

Ao estudar os antigos primário e ginásio, Jallade (1978) mostra que os grupos sócio-econômicos mais elevados não só possuem uma maior escolarização como recebem um retorno maior. A importância da qualidade da escola nesses resultados deve ter sido primordial. Behrman e Birdsall (1983) concluem, com base no Censo de 1970, que o retorno social de uma melhoria na qualidade do ensino deve ser muito maior

---

1 Ver as resenhas apresentadas por Simmons e Alexander (1978), Schiefelbein e Simmons (1981) e para o caso brasileiro, por Brandão et al. (1983).

que o retorno apenas de uma maior escolarização. Uma forte evidência de que os fatores que atuam do lado da oferta, isto é, disponibilidade de vagas e qualidade do ensino, exercem uma clara influência sobre a demanda pode ser encontrada em Birdsall (1983). Usando informações do Censo de 1970, Birdsall verifica que a inclusão das variáveis de oferta reduzem apreciavelmente a influência da renda paterna na escolaridade. A elasticidade da demanda com respeito à disponibilidade e qualidade da escola é grande nas zonas urbana e rural. Na zona urbana, o efeito dessas variáveis é maior para as famílias mais pobres; o contrário parece acontecer na zona rural.

## UM MODELO LOGÍSTICO PARA A ANÁLISE DA PROGRESSÃO ESCOLAR

Freqüentemente, os estudos sobre os determinantes da escolarização abordam essa questão de uma maneira metodologicamente deficiente. Por exemplo, uma maneira bastante usual de se estudar esses determinantes é através da especificação de uma função linear com a forma

$$Y_i = \beta_0 + \sum_k \beta_k X_{ik} \quad (1)$$

onde  $Y_i$  é o número de anos de escolaridade completos para cada indivíduo  $i$ , os  $X_{ik}$  referem-se aos  $K$  fatores, geralmente relacionados com a origem sócio-econômica, que influenciam essa escolaridade, e os coeficientes  $\beta_k$  indicam os efeitos desses fatores sobre a escolaridade individual (e.g. Blau e Duncan, 1967; Silva, 1981).

Um problema sério com esse tipo de abordagem é que, ao mensurar esses efeitos sobre a escolaridade em termos de um único coeficiente (médio) para toda a amostra, descarta-se a possibilidade de que esses coeficientes variem de uma forma sistemática e significativa ao longo dos diversos níveis de escolaridade. Em outras palavras, um fator importante a um certo nível de escolaridade pode não o ser em um outro nível.

Existem razões para supor que esses determinantes tenham um comportamento previsível, relacionado com os níveis de escolaridade. Por exemplo, pode-se esperar que o efeito da renda familiar sobre a permanência na escola aumente com o nível de escolaridade. Nestes casos, a equação (1) deixa de representar o fenômeno ainda que possa se contornar essa deficiência obtendo estimativas para diversas faixas etárias.

Uma forma mais conveniente para se analisar a determinação da escolaridade é conceptualizá-la em termos de uma seqüência de transições entre patamares de escolaridade. Dessa forma, podemos estudar os fatores determinantes da escolaridade do indivíduo, mensurando-os através de um conjunto de probabilidades condicionais de progressão escolar. Essas probabilidades indicam as chances de um indivíduo atingir um certo nível de escolaridade dado que ele completou o nível imediatamente anterior.

Ou seja, se sabemos até onde o indivíduo progrediu na escola, podemos deduzir todas as transições entre patamares de escolaridade que ele foi capaz de

completar e, portanto, construir uma seqüência de variáveis dicotômicas que expressam se aquela pessoa completou ou não um patamar qualquer de escolaridade. Evidentemente, ao analisarmos uma transição, só consideramos aqueles indivíduos que completaram o nível anterior, pois só esses poderiam completar essa transição. Portanto, para cada nível de escolaridade, os indivíduos aptos a fazer a progressão se constituem em uma sub-população, cujo tamanho diminui ao se passar dos níveis mais baixos para os mais altos de escolaridade.

A redução do tamanho das sub-populações se deve ao processo de seleção dos indivíduos segundo suas origens sociais, associado à progressão escolar. Essa auto-seleção poderia dificultar a comparação dos efeitos entre duas transições quaisquer ao modificar a distribuição marginal das variáveis. Porém, recentemente, Mare (1980) sugeriu que um modelo logístico fosse aplicado a cada transição, codificada como uma variável binária.

O modelo logístico assegura uma análise adequada das alterações nas desigualdades de oportunidades educacionais, livre da contaminação oriunda do próprio processo de seletividade intra-escolar pois, ao contrário de modelos como o da equação (1), as estimativas dos coeficientes não são influenciadas por mudanças nas distribuições marginais das variáveis (Bishop et al., 1975, p.9-15). Como observa Mare, "diferenças de efeito entre sub-populações resultam de diferenças genuínas de associação entre as variáveis mensuradas" (Mare, 1980, p. 297). Igualmente, os efeitos das variáveis independentes sobre a realização de uma transição não são influenciados pela proporção dos que a fazem, como seria o caso em modelos mais simples como o da equação (1). Finalmente, as probabilidades de transição são assintoticamente independentes entre si, o que possibilita que a equação (2) seguinte seja estimada para cada transição escolar que se queira (Feinberg, 1977; 1980)<sup>2</sup>.

As equações desse modelo têm a forma

$$\log_e \left( \frac{\theta_{ij}}{1 - \theta_{ij}} \right) = \beta_{j0} + \sum_k \beta_{jk} X_{ijk} \quad (2)$$

onde  $\theta_{ij}$  é a probabilidade de que o  $i$ -ésimo indivíduo complete a  $j$ -ésima transição escolar,  $X_{ijk}$  é o valor da  $k$ -ésima variável independente para aquele indivíduo, e os  $\beta_{jk}$  são parâmetros indicativos dos efeitos dessas variáveis sobre o logaritmo das chances dele completar a transição. A especificação do modelo logístico é adequada no sentido de que, dada a formulação da variável dependente, mudanças unitárias nas variáveis independentes geram mudanças razoavelmente uniformes nessa variável, o que não aconteceria, por exemplo, se a especificação fosse do tipo linear nas probabilidades.

<sup>2</sup> No nosso caso, o modelo proposto foi estimado por procedimentos de máxima verossimilhança para dados não agrupados, utilizando-se para isso o programa DREG incluído no pacote estatístico OSIRIS IV.

## DADOS

Neste trabalho utilizaremos os dados provenientes da Pesquisa Nacional por Amostragem Domiciliar (PNAD), levada a cabo pela Fundação IBGE em 1976. A escolha da PNAD-76 prendeu-se ao fato dela ser o único levantamento a nível nacional a conter informações relativas à origem familiar e sócio-econômica dos indivíduos. Constitui portanto um conjunto de dados privilegiado para o tipo de investigação aqui proposto. Restringimos as informações à população masculina de idade variando entre 20 e 64 anos, perfazendo uma amostra de 12.755 homens. Isto porque pode-se considerar a escolarização como um processo já concluído para a maioria da população aos 20 anos. Por outro lado, as diferenças existentes nesse processo entre homens e mulheres tornaram recomendável considerar isoladamente a população masculina.

Como o levantamento dos dados foi feito em 1976, essas observações cobrem portanto os nascimentos que ocorreram de 1912 a 1956, ou seja, quase toda a primeira metade do presente século. Agrupando os indivíduos em coortes de nascimento de cinco anos, obtém-se 9 coortes que nos mostram a evolução histórica da escolarização no Brasil. A Tabela 1 apresenta para cada coorte de nascimento a média, o desvio-padrão e o coeficiente de variação da variável anos de escolaridade dos indivíduos.

**Tabela 1**  
**Tendências da realização educacional**  
**Brasil (1976) – homens de 20 a 64 anos**

Anos de Nascimento da Coorte	Anos de Escolaridade da Coorte		
	Média	Desvio-Padrão	Coeficiente de Variação
1912-1916	2,237	3,208	1,379
1917-1921	2,517	3,336	1,326
1922-1926	3,095	3,801	1,228
1927-1931	3,279	3,802	1,160
1932-1936	3,877	4,158	1,072
1937-1941	4,125	4,208	1,020
1942-1946	4,477	4,344	0,907
1947-1951	4,935	4,436	0,899
1952-1956	5,027	3,904	0,777

Fonte: Silva (1981)

Durante o período em observação, verifica-se uma melhoria na realização educacional dos indivíduos. Entre a coorte mais velha e a coorte mais nova, o nível de escolaridade médio dos brasileiros mais que dobrou, passando de aproximadamente 2,3 anos para mais de 5 anos de escolaridade. A distribuição da escolaridade vem melhorando sistematicamente ao longo do tempo, pois o valor do coeficiente de variação na coorte mais jovem é quase a metade daquele na coorte mais velha. No entanto, apesar desse au-

mento na escolaridade geral e dessa melhoria relativa na distribuição das chances educacionais, o nível médio de escolarização ainda é muito baixo, e a sua distribuição bastante assimétrica.

A PNAD-76 foi obtida por uma mostra estratificada por conglomerados relacionada em múltiplos estágios<sup>3</sup>. Assim sendo, as suposições de amostra aleatória simples inerentes aos procedimentos analíticos que adotaremos a seguir não são satisfeitas. Em particular, nossas estimativas dos erros padrões terão provavelmente um viés para baixo (Kish e Frankel, 1974; Mare, 1980), e portanto as estatísticas relativas aos testes de significância devem a rigor ser consideradas para fins meramente descritivos. No entanto, como indica Mare (1980) referindo-se a resultados obtidos por Rao e Scott (1979), o viés nos erros padrões estimados é aparentemente reduzido.

As variáveis incluídas na análise são as seguintes:

- $X_1$  – *ocupação do pai*: esta variável foi mensurada em termos de uma escala de posição sócio-econômica das ocupações, desenvolvidas por um dos autores a partir de informações contidas no Censo Demográfico de 1970 (Silva, 1973);
- $X_2$  – *educação do pai*: mensurada em termos de anos completos de escolaridade;
- $X_3$  – *lugar de nascimento do respondente*: variável dicotômica codificada com a forma  $x_3 = 1$  se o respondente nasceu em área urbana e  $x_3 = 0$  no caso contrário. É um indicador do diferencial rural/urbano no acesso à escolarização, refletindo, portanto, basicamente fatores de oferta.
- $X_4$  – *migração*: variável dicotômica codificada com a forma  $x_4 = 1$  se o indivíduo atualmente mora em área diferente daquela em que nasceu e  $x_4 = 0$  no caso contrário. Dado que o fluxo migratório é predominantemente de destino urbano, esta variável deve representar basicamente a melhoria nas condições de acesso à escolarização para aqueles indivíduos que migraram;
- $X_5$  – *cor do respondente*: de acordo com várias evidências (Silva, 1981), podemos codificar a variável cor do respondente através de uma dicotomia Branco/Não-Branco sem fazer muita violência à realidade sócio-econômica dos diversos grupos de cor. Assim, essa variável foi codificada com a forma  $x_5 = 1$  se o indivíduo pertence ao grupo "Branco",  $x_5 = 0$  no caso contrário.

É importante lembrar que essas variáveis trazem algumas limitações à análise. Em primeiro lugar, ocupação do pai representa a situação familiar do respondente no momento em que ele entrou para o mercado de trabalho, sendo portanto um fato posterior à escolarização de alguns respondentes e anterior ao término da escolarização dos respondentes que começaram a trabalhar antes de abandonar a escola. É, dessa forma, um fator indicativo de circunstâncias eventualmente diferentes no ciclo da vida do indivíduo, dificultando a interpretação dos efeitos dessa variável entre diferentes transições. Por outro lado, as variáveis

<sup>3</sup> Para uma descrição dos procedimentos amostrais da PNAD-76 veja-se FIBGE(1982).

que dizem respeito às circunstâncias do nascimento do indivíduo (por exemplo, cor ou lugar de nascimento) ou mesmo que, como educação do pai, correspondem a uma fase do ciclo vital paterno, que normalmente antecede o processo de escolarização em estudo, são mais confiáveis.

Uma limitação mais séria, advém do fato de que os dados não incluem nenhuma variável mediando o efeito da origem sócio-econômica familiar sobre a realização educacional dos indivíduos. Como já mencionado, o trabalho de Sewell e Hauser (1975) estabeleceu com bastante sucesso uma lista de fatores intervenientes no processo de estratificação educacional. Em particular, mostrou que a capacidade cognitiva dos estudantes, suas aspirações ocupacionais e a influência dos chamados "outros significantes" não só dependem positivamente da sua origem social familiar como exercem efeitos positivos independentes sobre a realização escolar. Assim sendo a ausência dessas variáveis impede uma compreensão desses mecanismos intervenientes.

Por exemplo, é de se esperar que a influência das variáveis de origem social sobre a realização educacional diminua com o nível de escolaridade devido ao seu efeito atenuado através das variáveis intervenientes. Mare (1980) sugere que os efeitos mediados pela capacidade cognitiva diminuem porque nos níveis mais elevados de escolaridade a inteligência dos que têm uma origem social baixa se torna mais homogênea, ao serem eliminados preponderantemente os de menor capacidade cognitiva. Como só foi estimada a forma reduzida de um modelo estrutural, não é possível distinguir na alteração dos coeficientes obtidos em diferentes transições o que é devido ao efeito direto das variações de origem social e ao efeito mediado pelas variáveis intervenientes.

Também, deve-se reconhecer que os efeitos do lado da oferta escolar não são adequadamente captados pelo modelo. Num país como o Brasil, onde o número de vagas é um forte condicionante das possi-

bilidades de acesso à escola e onde mesmo a escola pública revela diferenciações patológicas na qualidade do ensino, as chances de realizar uma transição são sensivelmente influenciadas por estes fatores. Ainda assim, pode-se esperar que variáveis de origem social captem em parte os fatores do lado da oferta, dada a colinearidade entre ambas. Além disso, a variável lugar de nascimento apresenta provavelmente o meio em que se deu a socialização educacional e portanto capta as diferenças de acesso e permanência na escola entre as zonas urbana e rural.

A variável dependente *realização escolar do indivíduo* foi decomposta numa série de 8 transições condicionais que refletem os momentos centrais do processo de progressão escolar. São as seguintes as transições escolhidas: 1) completar a 1ª série; 2) completar o primário, dado que completou a 1ª série; 3) entrar para o ginásio, dado que completou o primário; 4) completar o ginásio, dado que entrou para o ginásio; 5) entrar para o colegial, dado que completou o ginásio; 6) completar o colegial, dado que entrou para o colegial; 7) entrar para a universidade, dado que completou o colegial; 8) completar a universidade, dado que entrou para a universidade.

A descrição dos ciclos educacionais segue a denominação antiga uma vez que ela corresponde à da experiência de quase a totalidade dos indivíduos na amostra.

## RESULTADOS

Um primeiro ponto a observar diz respeito ao atrito permanente nessa série de transições que representam o processo de escolarização. Dos 12.755 indivíduos que compõem a amostra e que portanto poderiam efetuar ou não a primeira transição, apenas 761 (5,9%) chegaram a se habilitar a passar pela última transição. O fracasso na 1ª série do primário é de 22% e a percentagem dos que embora realizassem a primeira transição não concluíram o primário é de 42%.

**Tabela 2**  
**Medidas descritivas das variáveis independentes em cada transição**

Nível de Escolaridade (transição)	Médias e Desvios-Padrões				Percentagens dos que completam a transição			
	Ocupação do Pai X	S <sub>x</sub>	Educação do Pai X	S <sub>x</sub>	Nascimento em Área Urbana	Migrante	Cor Branca	Indivíduos Completando a Transição
Completar a 1ª Série	12,6	9,8	1,9	2,8	46,0	55,7	56,8	77,8
Completar Primário	13,2	10,3	2,3	3,0	53,9	57,7	61,7	42,3
Entrar Para Ginásio	15,8	12,9	3,6	3,8	74,8	61,1	67,5	81,4
Completar Ginásio	16,5	13,7	4,0	4,0	79,3	62,0	69,8	66,9
Entrar para Colegial	18,4	15,2	4,7	4,3	82,5	64,3	73,7	73,4
Completar Colegial	19,9	16,7	5,3	4,6	84,1	65,3	75,7	79,2
Entrar Universidade	21,0	17,5	5,7	4,7	85,8	66,4	79,1	57,3
Completar Universidade	22,8	19,3	6,5	4,9	87,5	68,9	82,4	58,0

Fonte: PNAD-76, Tabulações Especiais

Mesmo entre os indivíduos que poderiam entrar para o nível universitário, apenas 57% completaram com êxito essa transição. Verifica-se portanto uma extrema redução no fluxo de alunos ao longo do processo de escolarização.

Anteriormente, vimos que o efeito das variáveis de origem social sobre as chances de progresso na escola é intermediado por um conjunto de variáveis psico-afetivas (Sewell e Hauser, 1975). Portanto, o efeito final da origem social sobre as chances de progresso depende da estrutura de covariância entre as variáveis de *background* e as características intervenientes. A seletividade escolar é influenciada pelas variáveis intervenientes, de tal forma que aqueles com maiores estímulos de seus pais, amigos e professores, maiores aspirações profissionais e maior capacidade cognitiva têm maiores chances de experimentar transições entre níveis mais elevados. Com isso, a covariância entre os fatores de origem social e características intervenientes tende a enfraquecer ao longo do processo educacional, diminuindo assim o efeito total da origem social dos alunos sobre suas chances de progresso escolar.

A Tabela 3 apresenta o resultado das regressões para o modelo logístico relativas às oito transições propostas. Em cada coluna da Tabela 3 são apresentadas a estimativa dos coeficientes das variáveis independentes e, entre parênteses, a razão entre esses coeficientes e seus erros padrões. Supondo-se uma amostra aleatória simples, essas razões se distribuem assintoticamente como uma curva normal padrão sob a hipótese nula de que os coeficientes são nulos. Assim, valores dessa razão superiores a 1,96 serão significantes ao nível de 0,05. Como estamos trabalhando com uma escala logarítmica, os valores dos coeficientes podem ser interpretados como o efeito de mudança em uma unidade da variável independente sobre a diferença percentual nas chances de sucesso na transição.

Na base de cada coluna, são registrados " $R^2$ ",  $X^2$  e N. O " $R^2$ " mede a proporção de melhoria no erro de predição em relação à hipótese nula de que todos os coeficientes são nulos (Dumochel, 1976). O  $X^2$  é a estatística usual de qui-quadrado que testa a qualidade de ajustamento do modelo e o N indica o tamanho da amostra para cada transição.

**Tabela 3**  
**Coefficientes dos efeitos dos fatores de origem social sobre a progressão escolar (a)**

Coeficiente	Transição							
	Completar 1ª Série	Completar Primário dado que completou 1ª série	Entrar para Ginásio dado que completou o Primário	Completar Ginásio dado que entrou para o Ginásio	Entrar para Colegial dado que completou o Ginásio	Completar Colegial dado que entrou para o Colegial	Entrar para Universidade dado que completou o Colegial	Completar Universidade dado que entrou para a Universidade
Ocupação do Pai (b)	0,0015 [4,48]	0,0029 [8,98]	0,0015 [2,83]	0,0026 [5,46]	0,0019 [3,59]	0,0012 [2,11]	0,0003 [0,82]	0,0016 [3,06]
Educação do Pai (b)	0,0638 [25,62]	0,0301 [23,07]	0,0168 [8,60]	0,0149 [10,17]	0,0118 [7,27]	0,0064 [3,54]	0,0073 [4,73]	0,0027 [1,42]
Nascimento em Área Urbana	0,6790 [22,57]	0,6043 [27,14]	0,2492 [11,15]	0,0844 [4,37]	0,0328 [1,49]	0,0702 [2,75]	0,0231 [0,98]	-0,0139 [-0,48]
Migrante	0,1758 [8,29]	0,1530 [7,63]	0,1575 [4,70]	0,1546 [5,07]	0,0770 [2,09]	0,1190 [2,62]	0,1107 [2,71]	0,0139 [0,27]
Cor Branca	0,2741 [13,22]	0,0869 [4,70]	0,1288 [4,54]	0,1107 [4,49]	0,0429 [1,51]	0,1659 [5,22]	0,0727 [2,50]	0,0300 [0,87]
Constante	0,6185 [13,26]	-1,3506 [-27,98]	0,8790 [10,36]	-0,1651 [-2,16]	0,2340 [2,49]	0,8692 [7,71]	-0,1751 [-1,78]	-0,2011 [-1,51]
" $R^2$ "	0,162	0,144	0,076	0,064	0,044	0,042	0,024	0,018
$X^2$	2.741,24	2.607,56	387,01	378,36	154,90	91,80	61,86	26,52
N	12.755	9.921	4.194	3.412	2.282	1.676	1.328	761

(a) entre parênteses estão as razões entre os coeficientes e seus erros-padrões

(b) as variáveis Ocupação do Pai e Educação do Pai têm seus valores multiplicados por 10

Fonte: IBGE, PNAD-76: **Tabulações Especiais**

Como podemos observar pelos valores de "R<sup>2</sup>", o poder explicativo das variáveis declina monotônica e rapidamente da primeira para a última transição escolar. Na primeira transição, compreendendo as chances de se completar a 1ª série, as variáveis reduzem em 16,2% o erro preditivo, enquanto que para as chances de se completar a terceira transição (entrar para o ginásio dado que completou o primário) o poder explicativo do modelo se reduz em mais da metade, diminuindo em apenas 7,6% o erro de predição. Quando chegamos à última transição, a influência das variáveis nas chances de completar com sucesso essa transição é quase nula, atingindo uns escassos 1,8% de redução no erro de predição. Dessa forma, as expectativas elaboradas anteriormente quanto ao comportamento geral das variáveis que, predominantemente, descrevem o efeito da origem social sobre a progressão escolar são confirmadas quando consideramos essas variáveis como um todo.

Os efeitos de cada variável independente também parecem declinar ao longo das transições, embora em alguns casos esses declínios não sejam monotônicos. Por exemplo, o fato de nascer em área urbana dá ao indivíduo uma vantagem da ordem de 68% a mais nas chances de completar a 1ª série; essa vantagem declina para 25% a mais na transição relativa à entrada no ginásio e virtualmente desaparece no ponto de entrada do ciclo universitário. Similarmente, uma diferença de educação paterna da ordem de um ano de escolaridade corresponde a uma diferença de 6,4% nas chances de completar a 1ª série, mas o coeficiente não é significativo para a última transição.

Os coeficientes das variáveis apresentam o sinal esperado em todas as transições, salvo nascimento em área urbana na última transição, quando deixa de ser significativo. A variável nascimento em área urbana tende a predominar sobre as demais até a terceira transição (entrar para o ginásio). Mais importante, é que o seu coeficiente tem uma significância equivalente à do coeficiente de educação paterna nessas transições. A importância dessa variável deve ser atribuída ao maior acesso ao sistema escolar na zona urbana, que deve estar representando. Ou seja, aparentemente, existem efeitos claros e fortes nas três transições iniciais de fatores representativos da oferta de escolaridade, especialmente nas chances de completar o primário.

Nas transições entre completar o ginásio e entrar para a universidade aparece com destaque a condição de migrante. Como o fluxo migratório predominante é de destino urbano e o fato de ser não migrante urbano já está representado pela variável nascimento em área urbana, as vantagens de ser migrante devem também estar associadas a uma maior facilidade no acesso e permanência na escola. De novo, condições do lado da oferta parecem predominar embora a busca de escolaridade seja uma das causas de migração. Educação paterna tem um efeito expressivo, ainda que, como vimos, declinante. Cor branca, representando efeitos de origem e discriminação racial, tem um elevado impacto exatamente na primeira transição.

A única variável a discrepar claramente da tendência de efeitos declinantes ao longo das transições escolares é a ocupação paterna. Aqui transparece outro padrão importante, qual seja, o fato desse efeito

ser mais elevado e significativo nas transições correspondentes ao fim dos diversos ciclos educacionais. Com exceção do colegial, o valor do coeficiente é mais alto para se completar o ciclo do que para entrar naquele mesmo ciclo. Como a variável ocupação do pai, frente à presença de uma outra variável mensurando o nível educacional e cultural paterno, tende a capturar os aspectos mais estritamente econômicos da situação paterna, esse padrão parece indicar a relevância da condição financeira do pai na continuação dos estudos. Em outras palavras, se para entrar num certo ciclo educacional a situação sócio-econômica familiar do aluno é importante determinante, suas chances de completar com sucesso esse ciclo parecem depender particularmente da renda familiar. Neste caso, esta variável poderia estar captando alguns efeitos da qualidade do ensino que não foi possível considerar explicitamente.

## CONCLUSÕES

Este trabalho utilizou um modelo logístico para estudar a influência das variáveis de origem social sobre as chances de progressão escolar. Este modelo apresenta vantagens sobre os seus congêneres, onde a variável dependente é ou anos de escolaridade ou a probabilidade de completar uma transição escolar, porque é capaz de medir os efeitos dos fatores de origem social independentemente das alterações nas distribuições marginais das variáveis em cada transição. Os dados disponíveis não permitiram a inclusão de variáveis intervenientes ou ligadas à oferta escolar. Ainda assim, em sua forma reduzida, o modelo conseguiu mostrar a importância das variáveis de origem social consideradas, destacando-se a educação paterna. Nas primeiras transições, sobressai a variável área de nascimento, que deve estar captando efeitos ligados à oferta de escolaridade. Como esperado, geralmente os coeficientes das variáveis diminuem monotonicamente à medida que as transições avançam. A exceção é o *status* ocupacional do pai, mais importante para completar um ciclo educacional do que para iniciá-lo. Isto deve estar refletindo a relevância da renda familiar nas chances de concluir um ciclo educacional. Considerando-se que os dados referem-se a todo o país e a 50 anos de sua história, o valor preditivo do modelo é satisfatório, devendo aumentar para análises regionais ou de coortes.

---

## REFERÊNCIAS BIBLIOGRÁFICAS

---

- BEHRMAN, J. R. & BIRDSALL, N. The quality of schooling: quantity alone is misleading. *American Economic Review*. 73:928-46, 1983.
- BIRDSALL, N. *Public inputs and child schooling in Brazil*. The World Bank, 1983. mimeo
- BISHOP, Y. et al. *Discrete multivariate analysis*. Cambridge, Mass. MIT Press, 1975.
- BLAU, P. & DUNCAN, O. D. *The american occupational structure*. New York, Wiley, 1967.
- BLAUG, M. The empirical status of human capital theory: a slightly jaundiced survey. *Journal of Economic Literature*. 14: 827-55, 1976.



- BRANDÃO, Z. et al. *Evasão e repetência no Brasil: a escola em questão*. Rio de Janeiro, Achiamé, 1983.
- CONLISK, J. Determinants of school enrollment and school performance. *Journal of Human Resources*. 4:140-57, 1969.
- DUMOUCHEL, W. H. *On the analogy between linear and log-linear regression*. University of Michigan, Dept. of Statistics, 1976. [Technical Report nº 67].
- DUNCAN, B. Education and social background. *American Journal of Sociology*. 72:363-72, 1967.
- FEINBERG, S. E. *The analysis of cross-classified categorical data*. Cambridge, Mass. MIT Press, 1977.
- FIBGE. *Metodologia da Pesquisa Nacional por Amostragem de Domicílios na Década de 1970*. Rio de Janeiro, 1982.
- GARNIER, M. & HOUT, M. Inequality of educational opportunity in France and in the United States. *Social Science Research*. 5:225-46, 1975.
- GOUVEIA, A. J. Origem social, escolaridade e ocupação. *Cadernos de Pesquisa*. (32):3-30, 1980.
- GUSTMAN, A. L. & PIDOT Jr., G. B. Interactions between educational spending and student enrollment. *Journal of Human Resources*. 8:3-23, 1973.
- HANSEN, D. O. Age variations in the formation of educational and occupational career goals of brazilian youth: a cross-cultural test of the Wisconsin model. *Comparative Education Review*. 21:37-50, 1977.
- HANSEN, R. & FEATHERMAN, D. Equality of schooling: trends and prospects. *Sociology of Education*. 49:99-120, 1976.
- HUSÉN, T. Strategies for educational equality. In: Education inequality and life chances. Paris, O.E.C.D., 1975, v.1.
- JALLADE, J.P. Educação de Primeiro Grau e desigualdade de renda no Brasil: perspectivas a longo prazo. *Pesquisa e Planejamento Econômico*. 8:561-98, 1978.
- JOHNSON, G. E. & STAFFORD, F. Social returns to quantity and quality of schooling. *Journal of Human Resources*. 8:140-55, 1973.
- KERCKHOFF, A. C. Stratification process and outcomes in England and the United States. *American Sociological Review*. 39:789-801, 1974.
- KISH, L. & FRANKEL, M. Interference from complex samples. *Journal of the Royal Statistical Society*. Ser. B. 36:1-37, 1974.
- LERNER, M. *Determinants of educational attainment in Brazil, 1960*. Ann Arbor, University Microfilm, 1971.
- MARE, R. Social background and school continuation decisions. *Journal of the American Statistical Association*, 75:295-305, 1980.
- MASTERS, S. H. The effect of family income on children's education: some finding on inequality opportunity. *Journal of Human Resources*. 4:158-175, 1969.
- RAO, J. & SCOTT, A. The analysis of categorical data from complex surveys II: chi-squared tests for independence. University of Aucland, 1979. mimeo.
- SCHIEFELBEIN, E. & SIMMONS, J. Determinantes del rendimiento escolar reseña de la investigación para los países en desarrollo. Ottawa, International Development Research Center, 1981.
- SCHMIDT, I.B. & MIRANDA, G. *Determinantes da escolarização no Brasil*. Rio de Janeiro, ECIEL, 1977.
- SEWEL, W. & HAUSER, R. *Education occupation and earnings: achievement in the early career*. New York, Academic Press, 1975.
- SILVA, N. V. Cor e o processo de realização sócio-econômica. *Dados*. 25:391-409, 1981.
- . *Posição social das ocupações*. IBGE Centro de Informática, 1973.
- SIMMONS, J. & ALEXANDER, L. The determinants of school achievements in developing countries: a review of the research. *Economic Development and Cultural Change*. 26:341-57, 1978.
- SOUZA, A. M. *Financiamento da educação e acesso à escola no Brasil*. Rio de Janeiro, IPEA, 1979.
- TURNER, R. H. Sponsored and contest mobility and the school system. *American Sociological Review*. 25:855-67, 1960.
- WACHTEL, P. The effect of school quality on achievement, attainment levels and lifetime earnings. *Explorations in Economic Research*. 2:502-36, 1975.