

*Philip R. Fletcher**
*Sérgio Costa Ribeiro***

1. INTRODUÇÃO

A função da estatística educacional é de fornecer informações fatuais sobre o estado e funcionamento do sistema de ensino. Até hoje, no Brasil, as informações sobre o sistema de ensino têm se concentrado na descrição meramente quantitativa, a saber:

- 1) estoque e fluxo de alunos;
- 2) estoque, qualificação, salário e rotatividade dos professores;
- 3) dinâmica da infra-estrutura da rede de escolas.

Acompanhados por informações demográficas e de custos, pode-se projetar a evolução do sistema em seus aspectos quantitativos e permitir a administração de seu patrimônio.

No entanto, para analisar o sistema de ensino é preciso monitorar sistematicamente as práticas pedagógicas das escolas e as competências cognitivas de seus estudantes para qualificar o significado dos processos formais de ensino. Estes aspectos não têm merecido atenção das autoridades governamentais.

Esta situação não é inevitável. Nos Estados Unidos vários estados avaliam o desempenho das escolas públicas e particulares mediante baterias de testes padronizados administradas anualmente há trinta anos. Além do mais, existem estudos da população em geral que fornecem um **baseline** do domínio cognitivo da população adulta em diferentes áreas do país e em diferentes segmentos da sociedade.

O pouco trabalho que existe nessa área no Brasil são estudos localizados empreendidos por pesquisadores da área de educação. Motivados por preocupações acadêmicas e profissionais, sem nenhum planejamento por localizados empreendidos por pesquisadores da área de educação. Motivados por preocupações acadêmicas e profissionais, sem nenhum planejamento por

* IPLAN/IPEA-OIT-PNUD

** LNCC/CNPq

parte das autoridades governamentais, esses estudos não tem procedimentos metodológicos padronizados que permitam a comparabilidade de seus resultados e sua aplicação para qualificar e monitorar o desempenho do sistema.

A ênfase dada aos aspectos quantitativos da rede de ensino no Brasil é justificada considerando a insuficiência de matrículas no passado. Hoje quando existem matrículas quase suficientes para permitir que todas as pessoas de uma geração completem a oitava série do 10º grau, a questão da competência cognitiva dos alunos se torna cada vez mais importante. Nestas circunstâncias, uma preocupação exclusiva com o desempenho quantitativo se torna insuficiente para administrar o ensino.

A expansão desordenada do sistema de ensino evoluiu de uma maneira completamente caótica. Sem administração e a devida aplicação de controles pela sociedade, esta expansão provoca uma variação interna no sistema de tal ordem que o significado dos diplomas emitidos pelas escolas deixam de serem referencial de domínio cognitivo. Não se sabe o que se aprende nas escolas do País e, por conseguinte, o que a população efetivamente domina. Suspeita-se que a promoção a quinta série do 1º grau, num bairro de classe média na cidade de São Paulo, não tem, nem de longe, o mesmo significado desta promoção em Barra do Corda no Estado do Maranhão.

Avaliações cognitivas somativas padronizadas e sistemáticas tem que ser urgentemente implantadas por um órgão especializado. A implantação desse sistema teria que passar necessariamente por uma fase experimental a base de amostragem e, para ser bem sucedida, seria preciso ter um conhecimento bastante seguro dos índices de repetência, promoção e evasão existentes no País para saber onde implantar estas experiências.

No entanto, nem a parte quantitativa da estatística educacional está sendo desenvolvida atualmente com o devido rigor pelas duas unidades formalmente encarregadas de executar este serviço, o SEEC/ MEC e o IBGE. Existem divergências inaceitáveis entre essas duas fontes nas estimativas das matrículas, na população em idade escolar fora da escola e nas taxas de repetência, promoção e evasão.

Por exemplo, o SEEC acusa uma matrícula na primeira série do 1º grau de 5,8 milhões em 1970, enquanto o IBGE constatou apenas 4,7 milhões na ocasião do VIII Recenseamento Geral do mesmo ano, uma diferença de 19% a menos. Em 1980, as mesmas fontes acusaram 7,0 e (ainda) 4,7 milhões, uma diferença de 34%! Os dados do SEEC sustentam um crescimento médio anual nas matrículas da primeira série de 1,9% durante a década, enquanto os dados do IBGE mostram um crescimento de apenas

0,09% por ano. Nas áreas rurais, segundo o SEEC, as mesmas matrículas cresceram 2,1 % ao ano, enquanto o IBGE revela um *decrécimo* de 3,2% ao ano!

O mesmo tipo de discrepância, embora de menor magnitude, é encontrado quando só comparados os dados do SEEC de 1982 com os resultados da PNAD do mesmo ano. Além do mais, retroprojetos da PNAD de 1982 para o Censo de 1980 revelam ainda outras inconsistências entre estas duas fontes do IBGE.

Segundo a PNAD de 1982, 90% da população em cada geração tem acesso à escola, sem tendência de aumento nas taxas de participação no decorrer do tempo. De acordo com os dados do Censo de 1980, apenas 84% em cada geração tem acesso à escola através da primeira série, outra vez, sem tendência de aumento. De acordo com a PNAD, o acesso é quase universal; segundo o Censo, o acesso ainda apresenta aprecieis e persistentes deficiências.

Sobre a população em idade escolar fora da escola, o MEC acusa um número variável de pessoas mas que pode atingir 8,5 milhões. A PNAD de 1982, no entanto, estima este numero em torno de 5,3 milhões, o que revela uma discrepância de 40% a menos. O debate publico que se trava em torno deste problema carece de solidas bases para estimar os custos de uma eventual solução.

A respeito da dinâmica interna do sistema escolar, as diversas bases de dados e metodologias empregadas produzem resultados pouco coincidentes. As Tabelas 1 e 2 apresentam as taxas de repetência, promoção e evasão em 1982 segundo o SEEC e o IBGE.

As taxas de transição de série derivadas dos dados do SEEC se baseiam nas declarações da administração das escolas. A "evasão" registrada numa escola não representa necessariamente o abandono do sistema de ensino. Portanto, alunos que aparecem como "novos" numa segunda escola podem ser, na realidade, os "evadidos" da primeira, provocando uma dupla contagem que distorce a percepção do funcionamento do sistema como um todo. Esta distorção de origem metodológica é de tal ordem que o numero de alunos tidos como "novos" na primeira série do 1º grau ultrapassa em muito o tamanho de uma geração, superando assim as possibilidades demográficas!

Como a reprovação é muitas vezes responsável pela transferência de escola, muitos alunos "novos" na primeira série são, de fato, repetentes. Portanto, além de superestimar a incidência, distorcendo profundamente nossa percepção da eficiência do sistema. Se, por um lado, a evasão sugere uma falta de lugares nas escolas suficientes para acomodar todos os alunos, altos índices de repetência apontam sérias deficiências

TABELA 1

REPETÊNCIA, PROMOÇÃO E EVASÃO,
SEGUNDO OS DADOS DA PNADS-82
TAXAS DE TRANSIÇÃO

Série em 19981	Série em 1982								Prom.	Evad.	Mart.
	1	2	3	4	5	6	7	8			
1	0.338	0.610								0.051	1.000
2		0.151	0.778							0.071	1.000
3			0.116	0.778						0.106	1.000
4				0.093	0.721					0.185	1.000
5					0.174	0.678				0.148	1.000
6						0.129	0.743			0.128	1.000
7							0.105	0.761		0.135	1.000
8								0.088	0.667	0.245	1.000

TABELA 2

REPETÊNCIA, PROMOÇÃO E EVASÃO,
SEGUNDO OS DADOS DA SEEC/MEC
TAXAS DE TRANSIÇÃO

Série em 19981	Série em 1982								Prom.	Evad.	Mart.
	1	2	3	4	5	6	7	8			
1	0.296	0.449								0.255	1.000
2		0.207	0.703							0.090	1.000
3			0.169	0.738						0.093	1.000
4				0.134	0.818					0.048	1.000
5					0.227	0.634				0.138	1.000
6						0.199	0.700			0.102	1.000
7							0.170	0.730		0.100	1.000
8								0.123	0.764	0.114	1.000

TABELA 3

REPETÊNCIA, PROMOÇÃO E EVASÃO,
SEGUNDO AS ESTIMATIVAS DO MODELO PROFLUXO
TAXAS DE TRANSIÇÃO

Série em 19981	Série em 1982								Prom.	Evad.	Mart.
	1	2	3	4	5	6	7	8			
1	0.547	0.433								0.019	1.000
2		0.360	0.598							0.041	1.000
3			0.283	0.648						0.069	1.000
4				0.197	0.618					0.185	1.000
5					0.282	0.628				0.089	1.000
6						0.148	0.756			0.095	1.000
7							0.133	0.759		0.108	1.000
8								0.089	0.680	0.231	1.000

no processo ensino-aprendizagem e levam à conclusão oposta: o sistema tolera e incentiva a retenção dos alunos. Grande parte do esforço para manter o aluno na escola redundava na repetência.

Os dados do IBGE, provenientes do PNAD 1.02 - Questionário de Educação - 1982, baseiam-se nas declarações dos pais e responsáveis dos alunos, que deveriam declarar a série cursada no ano letivo anterior. Como a repetência pode ser considerada um estigma social, existe um incentivo para o declarante evitar que o aluno apareça freqüentando a mesma série durante dois anos consecutivos. Escondendo o fato que o aluno já foi reprovado na série, suspeita-se que o fenômeno da repetência é grosseiramente subestimado na estatística baseada em declarações de pais e responsáveis.

Um modelo matemático do fluxo dos alunos no ensino de 1º grau, que não depende de declarações sobre o passado escolar das pessoas e depende tão-somente da situação escolar no momento da coleta dos dados, revela taxas de transição radicalmente diferentes das duas fontes consideradas acima, sobretudo, nas primeiras séries do 1º grau. Baseada no modelo PROFLEXO aplicado às declarações do PNAD 1.01 - Questionário de mão-de-obra - 1982, a Tabela 3 demonstra que mais da metade dos alunos na primeira série são repetentes e que apenas uma parcela insignificante da matrícula abandona a escola após a primeira série. O modelo revela que o número de matrículas de primeira série é mais de duas vezes o necessário para atender a demanda da população escolar!

Constatamos, portanto, que as diferentes metodologias levam a conclusões desbaratadas, o que levanta sérias dúvidas a respeito da qualidade das estatísticas educacionais. A situação é especialmente lamentável, considerando que estas são praticamente as únicas informações sistemáticas disponíveis no Brasil para avaliar o desempenho das escolas e dos alunos.

2. BREVE HISTÓRICO DA PNAD

Na década de 60, o êxito da experiência de países como os Estados Unidos e o Canadá em pesquisas contínuas de população veio ao encontro dos anseios da Aliança para o Progresso quanto a levantamentos que proporcionassem informações para o planejamento do desenvolvimento sócio-econômico.

A United States Agency for International Development (USAID) e o Inter-American Statistical Institute (IASI) trabalharam no sentido de levar os

países latino-americanos a desenvolver e implantar um sistema de inquéritos contínuos domiciliares para fornecer informações econômicas e sociais de interesse de cada país, permitindo ainda comparações entre os diversos países da região.⁽¹⁾ No Brasil, a PNAD se desenvolveu dentro desta concepção, adaptando as idéias contidas no chamado "Plano Atlântica", que serviu como ponto de partida para a elaboração da pesquisa.

O Plano Atlântica foi elaborado pelo U.S.Department of Commerce, Bureau of the Census, incorporando sua própria experiência, as recomendações da Organização das Nações Unidas e os acordos inter-americanos celebrados na véspera dos Recenseamentos 1960. Este plano, apresentado no México em 1962, foi divulgado em 14 volumes, cada um compreendendo uma etapa do planejamento e execução das pesquisas contínuas de base domiciliar. Ia até aos mínimos detalhes dos objetivos, conteúdo e layout (e cor de papel!) dos questionários, plano de amostragem, princípios de expansão da amostra, organização do trabalho de campo, manuais de entrevistadores, distribuição e recolhimento do material de pesquisa, crítica e codificação, plano tabular, análise dos dados e ainda as metodologias de projeções populacionais.⁽²⁾

Apesar de ser um trabalho extremamente ambicioso na abrangência e detalhe de suas propostas técnicas, percebe-se a intenção de estabelecer uma série de investigações bastante "leves", tanto em termos do conteúdo temático abordado nos questionários como no tamanho da amostra. O plano original era ambicioso apenas quanto a sua periodicidade trimestral, proposta para acompanhar a evolução da conjuntura econômica.

A duração das entrevistas deveria ficar entre 15 e 20 minutos em cada domicílio. A amostra prevista consistiria de apenas 10.000 domicílios distribuídos pelo País inteiro. Argumentava-se que uma amostra deste tamanho proporcionaria estimativas confiáveis num determinado momento e ainda seria útil para detectar mudanças de ano a ano. A periodicidade contemplada previa quatro pesquisas por ano, uma em cada trimestre. O trabalho de campo se desenvolveria durante o ano todo.

Inerente ao projeto das pesquisas contínuas, em cada período de inquérito, além de uma estrutura de informações contínuas sobre características demográficas básicas e emprego, um determinado terra adicional seria abordado, em caráter suplementar ou especial, com temática e periodicidade pré-estabelecidas. A periodicidade dos levantamentos de cada terra suplementar ou especial seria determinada de acordo com a natureza de sua variação temporal.⁽³⁾

As PNADs foram concebidas com esta filosofia de "criar um sistema relativamente fácil de manter, que investigasse algumas dezenas de milhares

de domicílios de cada vez, permitindo, a nível agregado, [...] um acompanhamento sistemático do quadro sócio - econômico do País".⁽⁴⁾ Os resultados iniciais da PNAD "vieram consagrar a aplicação da Técnica de Amostragem, como processo [...] para obtenção de dados considerados essenciais para o desenvolvimento dos estudos básicos para estabelecimento dos planos e programas que deverão acelerar o desenvolvimento econômico e social do País!"⁽⁵⁾

Mesmo assim, ingerências provenientes principalmente de órgãos de planejamento regional (SUDAM, SUDECO, SUDENE e SUDESUL), habituados com a apresentação de dados censitários com detalhamento até o nível municipal, queriam resultados das PNADs representativos, pelo menos, a nível de Unidade da Federação. O plano de amostragem original, de 30.000 domicílios, foi ampliado com o objetivo de atender esta demanda e ainda fornecer informações representativas das nove Regiões Metropolitanas. Em 1982, a amostra já abrangia 110.000 domicílios e 550.000 pessoas.

As PNADs foram implantadas no segundo trimestre de 1967 com periodicidade trimestral. Interrompida em 1970 por causa do VIII Recenseamento Geral, a série foi reiniciada no quarto trimestre de 1971 com periodicidade anual. Em 1973, foi novamente interrompida nos dois anos seguintes por causa do ENDEF. Em 1976, a série foi mais uma vez reiniciada, estendendo-se até o presente, numa base anual, interrompida apenas na ocasião do IX Recenseamento Geral de 1980.

Nem o conteúdo temático das pesquisas suplementares e especiais foi planejado de antemão e nem se conhece sua periodicidade prevista. A elaboração de seu conteúdo prossegue em consulta aos ministérios, técnicos do IPEA, professores universitários e pessoal dos próprios quadros do IBGE, geralmente em regimes de urgência e circunstâncias onde a tomada de decisões e a responsabilidade de cada um nem sempre são transparentes. Estas contingências levaram o IBGE a adotar uma amostra gigantesca para a PNAD, uma periodicidade anual e uma aparente arbitrariedade em sua escolha temática. Em anos recentes, o preparo dos arquivos de dados da pesquisa têm apresentado atrasos que prejudicam a atualidade das informações e sua utilidade em qualquer projeto de acompanhamento ou planejamento social.

3. ALGUNS PROBLEMAS METODOLÓGICOS DAS PNADs

A metodologia adotada pelo IBGE para as PNADs incorpora um "plano de amostragem [que] não foge ao esquema clássico de desenhos autoponderados, em estágios múltiplos, popularizados na década de 60 para

uso em pesquisas domiciliares".⁽⁶⁾ Este esquema, baseado na seleção dos elementos da amostra com probabilidades proporcionais ao tamanho, permite um processo de expansão extremamente simples e o cálculo direto dos erros das estimativas a partir da amostra.

No entanto, estas vantagens não são aproveitadas nas PNADs, devido ao método particular de expansão da amostra empregado. A expansão dos resultados da amostra não é feita mediante o uso dos estimadores naturais resultantes do plano amostral, mas sim através de projeções demográficas elaboradas pelos técnicos do IBGE.

No caso das PNADs da década de 80, essas projeções foram elaboradas a partir dos dados dos Recenseamentos Gerais de 1970 e 1980 e sob hipóteses associadas a taxas específicas de fecundidade, de mortalidade e migração. Seu grau de precisão está intimamente ligado ao das hipóteses feitas para essas taxas e, naturalmente, quanto mais distantes as projeções estiverem dos anos bases, menos válidas são estas hipóteses.⁽⁷⁾

O esquema clássico não incorpora um modelo de crescimento demográfico. Ao contrário, são as amostras que deveriam revelar o crescimento da população. Neste particular, as PNADs fogem do esquema proposto originalmente. Postula-se um "esquema clássico de desenhos autoponderados" e, logo em seguida, joga-se fora os potenciais benefícios deste plano, misturando erros amostrais, erros de recenseamentos anteriores e erros de um modelo matemático elaborado a partir dos recenseamentos. Nestas circunstâncias, quais seriam os erros reais dos parâmetros populacionais estimados a partir da PNAD?

Desde 1977, nos volumes de divulgação das PNADs, o IBGE fornece tabelas contendo os coeficientes de variação para ordens de grandeza de valores selecionados. No entanto, estas tabelas não incorporam o erro proveniente do modelo de crescimento demográfico utilizado na expansão da amostra.

Para estimar o erro de amostragem proveniente da seleção dos elementos da amostra, o usuário procura encontrar nessas tabelas a região e o tipo de variável de seu interesse. Em seguida, tenta identificar a linha da tabela cuja ordem de grandeza mais se parece com o número de pessoas encontradas com uma determinada característica. Lê na coluna apropriada o coeficiente de variação e assim calcula o erro padrão da estimativa.

Este procedimento é trabalhoso, impreciso e inapropriado. Os resultados das próprias PNADs revelam coeficientes de variação que variam entre si por um fator cinco ou mais.⁽⁸⁾ Esses coeficientes dependem do *design effect* da amostra, que varia de acordo com a variável e o conjunto de pessoas em consideração. Assim, para o País como um todo, uma variável como "pessoas

ocupadas em atividades "não-agrícolas" tem um coeficiente de variação 1,0%, enquanto "pessoas ocupadas em atividades agrícolas com rendimentos mensais de mais de 5 salários mínimos" tem um coeficiente de variação de 4,8%.⁽⁹⁾

A variabilidade desses coeficientes reflete largamente a correlação intra-classe da variável em consideração, o que representa, em última análise, o grau de segregação residencial da população segundo a variável. Outra variável qualquer, não relacionada com a segregação habitual como, por exemplo, "mulheres de 30 a 40 anos de idade", teria um coeficiente de variação bem menor. Nestas circunstâncias, é muito difícil para o usuário descobrir qual seria o erro padrão real de suas estimativas.

A identificação, no arquivo de dados, dos estratos e aglomerados (*clusters*) estabelecidos no momento da seleção da amostra são essenciais para o cálculo dos erros das estimativas dos parâmetros populacionais. Como estas informações não estão contidas nos arquivos de uso público, os erros não podem ser calculados pelos usuários externos. Portanto, os arquivos fornecidos pelo IBGE não são mensuráveis, quer dizer, a inferência de valores populacionais não pode se basear em medidas de variação, normalmente erros padrão, calculadas a partir dos dados da amostra.

Isto constitui um erro metodológico primário e injustificável que limita profundamente a utilidade das PNADs. E desprezar todos os benefícios de amostragem probabilística e a teoria estatística. Não se dispõe de critérios adequados para avaliar se uma estimativa gerada pela amostra reflete uma característica populacional: o erro da estimativa pode ser até maior que a própria estimativa! Nestas circunstâncias, o usuário não pode fornecer, com segurança, as estimativas dos erros amostrais e, portanto, faltam critérios adequados para julgar seus resultados.

Em sociologia existe um quase paradigma: as características de uma sociedade estão intimamente ligadas ao nível sócio-econômico. No Brasil, que apresenta um dos maiores índices de Gini em sua distribuição de renda, esta co-variação deve ser ainda mais marcante.

Nestas circunstâncias, estranha-se o fato que os indicadores de nível sócio-econômico não são utilizados para efeito de estratificação das amostras da PNAD. A falta desse controle pode ocasionar vieses sócio-econômicos na amostra. Seria perfeitamente viável evitar eventuais vieses dessa natureza se fosse feita uma estratificação múltipla, tanto geográfica como por nível sócio - econômico, na hora da seleção da amostra.

Isto seria também coerente com o propósito de apresentar indicadores contínuos sobre mão-de-obra, emprego e renda. No caso dos demais

indicadores sociais, por exemplo, os do suplemento de educação, a apresentação dos resultados segundo nível sócio-econômico deveria receber maior ênfase para mostrar as circunstâncias da condição de pobreza e as falhas na satisfação das necessidades básicas da população que deveriam ser as preocupações maiores do cidadão e dos poderes públicos. A estratificação da amostra segundo níveis sócio-econômicos seria, neste caso, consistente com os objetivos de investigação.

Finalmente, a metodologia das PNADs não prevê a aplicação da técnica de teste-reteste aos quesitos, o que impede uma avaliação de sua fidedignidade. "Em 1978, foi aplicada uma reentrevista a uma subamostra dos domicílios investigados [na PNAD], a fim de se obter uma idéia da variabilidade das respostas. Entretanto, este material até [1979] não foi analisado de modo a fornecer subsídios sobre o assunto".⁽¹⁰⁾ Seria aconselhável analisar a fidedignidade dos itens dos questionários da PNAD a cada ano e publicar esta informação junto com o plano tabular da pesquisa.

Evidentemente, estes comentários se aplicam a todos os temas investigados nas PNADs e, de maneira alguma, restringem-se apenas ao suplemento de educação de 1982.

4. O SUPLEMENTO DE EDUCAÇÃO DE 1982

No Anexo 1, estão listados os temas do PNAD 1.02 - Questionário de Educação - 1982, que mostram preocupações com:

- 1) indicadores do nível sociocultural das pessoas, com existência de livros, instrução e ocupação da família e de seus ascendentes;
- 2) informações sobre a situação escolar dos alunos: se a escola é pública ou particular, as taxas pagas a escola, o material didático e participação na merenda escolar;
- 3) estimativa do tempo de permanência no trabalho e na escola, do percurso casa/escola, dos afazeres domésticos, do hábito de assistir televisão, quantas vezes por semana as pessoas praticam alguma atividade física e quantos alunos executam trabalhos escolares feitos em casa;
- 4) informações sobre a situação de ex-alunos: se a escolarização foi pública ou particular, o motivo porque deixou de freqüentar a escola, se pretende voltar a escola no ano que vem e a idade em que concluiu seus estudos;

- 5) informações sobre a alfabetização e regressão ao analfabetismo;
- 6) subsídios para avaliar a repetência, promoção e evasão da matrícula de um ano para outro;
- 7) informações sobre a educação supletiva de 1º e 2º graus.

Junto com as informações do PNAD 1.01 - Questionário de Mão-de-Obra, pode-se relacionar a esfera escolar com o mundo do trabalho, a partir dos 7 anos de idade.

Vale a pena mencionar a complexidade da seqüência lógica dos blocos e dos quesitos no Questionário de Educação, que exige muita atenção dos entrevistadores e dos usuários dos dados (veja Anexo 1).

A divisão de idades entre 0 e 6 e de 7 anos em diante para informações sobre pré-escola e cursos regulares de 1º grau são excedentes entre si. Por conseguinte, fica impossível avaliar certos aspectos do pré-escolar. Se houver pessoas além dos 7 anos de idade no pré-escolar, não serão levantadas pelo Questionário.

A PNAD de 1982 não oferece possibilidades para relacionar a experiência pré-escolar com o desempenho da população no ensino de 1º grau. Seria de interesse perguntar quantos alunos beneficiaram-se da pré-escola na sua infância. A omissão de um quesito recordatório sobre esta experiência é lamentável. Sem estas informações, não podemos verificar se a experiência pré-escolar contribui para reduzir as altas taxas de repetência nas primeiras séries do 1º grau.

Questionamos o limite inferior de 7 anos para a aplicação dos itens referentes ao curso regular de 1º grau no PNAD 1.02 - Questionário de Educação - 1982, já que através do PNAD escolarização a partir dos 5 anos de idade, verificamos que 53% das pessoas que irão a escola (a maioria, alias!) já iniciaram seus estudos de 1º grau antes dos 7 anos. ⁽¹¹⁾

Por conseguinte, o Questionário de Educação exclui uma parcela apreciável da matrícula da primeira série do 1º grau. É irônico observar que, pelo menos neste aspecto, o Questionário de Mão-de-obra é melhor em matéria de educação que o próprio Questionário de Educação!

Todos os dois questionários da PNAD de 1982 deixam implícitas duas questões de fundamental importância:

- 1) Freqüente escola atualmente?
- 2) Já freqüentou escola alguma vez?

A primeira pergunta é implícita na seqüência de quesitos sobre freqüência a escola, mas é, também, crítica para o preenchimento de todos os blocos do questionário a seguir. Não nos parece uma questão para deixar a interpretação dos entrevistadores; a pergunta e suas opções de resposta

devem aparecer explicitamente no Questionário.

A segunda visa resolver uma ambigüidade comum a todas as pesquisas do IBGE: não existe nenhuma distinção entre as pessoas que nunca freqüentaram escola e as que freqüentaram mas não conseguiram concluir com êxito a primeira série do 1º grau. Esta controvérsia está relacionada com a polêmica questão do acesso a escola.

Propomos o seguinte esquema para o PNAD 1.01 - Questionário de Mão-de-Obra para resolver os dois problemas:

Nº Orde m (1)	-----	PARA PESSOAS DE 5 ANOS OU MAIS						
		Saber ler e escrever	Freqüência escolar	Curso que freqüenta		Freqüentou escola no passado	Última série concluída	
				Série	Grau (encerre)		Série	Grau
	-----	2- Sim 4 - Não	1- Sim (-)18 3 - Não (-) 20			2- Sim (-) 21 4 - Não (encerre)		
	-----	(16)	(17)	(18)	(19)	(20)	(21)	(22)
1	-----	⌋	⌋	⌋	⌋	⌋	⌋	⌋
02	-----	⌋	⌋	⌋	⌋	⌋	⌋	⌋
03	-----	⌋	⌋	⌋	⌋	⌋	⌋	⌋

A parte do PND 1.02 - Questionário de Educação - 1982 referente a freqüência escolar contém três perguntas que tentam sondar motivos compartimentais com alternativas de resposta pouco excedentes. A experiência de questionários fechados demonstra que perguntas que começam com a expressão "POR QUE..." tem pouca fidedignidade. Normalmente, os motivos para uma decisão são múltiplos, nem sempre conscientes na hora da decisão e muito menos vários anos depois. Nem sempre o informante, parente ou responsável compreende o comportamento dos outros membros do domicílio e muito menos seus motivos.

Os quesitos sobre a alfabetização, os únicos que tratam do domínio cognitivo na PNAD, também, têm pouca confiabilidade, na medida em que suas respostas dependem exclusivamente da opinião do entrevistado, sem a devida comprovação dessa capacidade na hora da entrevista. Estas perguntas podem até ter uma certa fidedignidade, no sentido de sua reprodutibilidade, mas carecem de validade como medidas de domínio cognitivo.

Em análises de dados censitários levantados com a mesma metodologia

da PNAD, verificamos que a pretensa aquisição da alfabetização (segundo critérios censitários) é simultânea com o ingresso do aluno na escola de 1º grau em muitas áreas do País. Parece que, em vários municípios do Nordeste brasileiro, a noção de "saber ler e escrever" coincide exatamente com a simples matrícula na escola. Qualquer avaliação de domínio cognitivo requer uma metodologia e um instrumental bem mais elaborados que se enquadrariam num inquérito especial, a exemplo do ENDEF na área de orçamentos familiares e nutrição.

Como já mencionamos na introdução, a tentativa de estimar as taxas de repetência, promoção e evasão na PNAD, através do registro da série cursada no ano letivo anterior, levou a valores inconsistentes com outras fontes e metodologias.⁽¹²⁾ Informações que comprometem a estima das pessoas são arriscadas. Seria preferível obter esse tipo de indicador através de modelos matemáticos do fluxo escolar, baseados em perguntas, as mais neutras possíveis, que não comprometam as pessoas.

O objetivo de descobrir até que ponto a necessidade de trabalhar afeta de maneira adversa a freqüência escolar transparece entre as preocupações da PNAD-1982. Infelizmente, as informações disponíveis não permitem inferir até que ponto os menores abandonam a escola porque foram forçados a trabalhar ou que trabalham porque foram forçados a abandonar a escola. No entanto, nem por aí resolve-se esta questão, porque o trabalho fornece condições necessárias que permitem ainda outros a continuarem seus estudos.

Que muitos alunos abandonam a escola por necessidade de trabalhar parece mais do que óbvio. Menos conhecido é o fato que a grande maioria dos alunos na faixa da obrigatoriedade escolar estudam em escolas que não oferecem instrução além da quarta série do 1º grau (informações, alias, que não são disponíveis na PNAD-1982). Suspeita-se, portanto, que muitos alunos "se evadem" do sistema escolar simplesmente porque não tem outra alternativa!

Em ainda outros casos, entre uma escola que oferece benefícios sociais precários e um trabalho que pelo menos oferece alguma remuneração, mesmo que modesta, muitos jovens vão dar preferência a segunda opção. Nestas circunstâncias, uma decisão com respeito ao trabalho pode ser implicitamente um juízo, negativo da utilidade da escola. Se a escolarização parece valer pouco ou impor um sacrifício - econômico ou psicológico - o trabalho surge como opção de maior relevância.

Evidentemente, a questão do menor e o mercado de trabalho no Brasil é indiscutivelmente conjugada à questão do ensino. Mas decisões com respeito ao trabalho não são necessariamente independentes de decisões

relativas a escola e vice-versa. O problema, do ponto de vista analítico, é que a PNAD-82 não oferece condições que permitam julgar a importância relativa dessas duas tendências.

As principais vantagens da PNAD são comuns a qualquer pesquisa de base domiciliar. Com visitas a domicílios, a pesquisa levanta informações sobre a população fora de escola, um assunto que não pode ser abordado em entrevistas conduzidas nas escolas, a exemplo dos dados levantados pelo SEEC e muitos estudos isolados de menor porte desenvolvidos por profissionais que trabalham na área de educação e escolham a escola como base para seus levantamentos.

Pesquisas do tipo PNAD podem ser utilizadas para mostrar as taxas de participação em cada série e grau do sistema escolar, quer dizer, a proporção em cada geração que atinge um determinado patamar de instrução. Por exemplo, os dados da PNAD mostram que 90% em cada geração ingressam na primeira série do 1º grau; 86% entram na segunda; e 73% chegam à quarta série. Mas apenas 56% em cada geração (62% da população escolar) ingressam na quinta série e apenas 38% de uma coorte de idade (42% da população escolar) atingem a oitava série do 1º grau. Este tipo de informação é útil para contestar a afirmação, bastante difundida nos meios de comunicação e no discurso público, que alega a evasão da metade dos alunos de 1º grau logo após a primeira série.

É um pouco mais trabalhoso calcular as taxas de repetência, promoção e evasão a partir dos quesitos sobre grau e série freqüentados, mas existem bases matemáticas bastante sólidas para esses círculos. A mesma fonte pode servir para estudar a idade de ingresso e aprovação em cada série, o que mostra sem margem de dúvida que os evadidos do sistema são os repetentes dos anos anteriores.

Estas informações fornecem valiosos subsídios para explicar porque as matrículas nas primeiras séries do 1º grau excedem em número o tamanho da população de referência e porque todo o esforço feito até hoje não conseguiu universalizar as oito séries de ensino compulsório contempladas na Constituição Brasileira.

O PNAD - Questionário de Educação - 1982 aborda diversos temas que poderiam ser explorados para aprofundar essas questões. Será que os repetentes recebem o mesmo número de horas de instrução que os outros alunos? Será que eles têm o mesmo acesso a material didático? Será que eles estudam em escolas que exigem trabalhos escolares feitos em casa? Será que as suas obrigações extra-escolares pesam desproporcionalmente em suas atividades diárias? Será que suas escolas são convenientemente localizadas e de fácil acesso? Estas são questões que o censo demográfico

não esclarece.

Pesquisas de base domiciliar permitem que estas questões sejam abordados em relação ao nível sócio-econômico das pessoas. Isto permite que o nível sócio-econômico seja analisado como determinante direto do desempenho dos alunos, como fator exógeno aos processos internos da escola ou ainda como um fator em interação. A análise da população de baixa renda mostra quanto ainda resta fazer para melhorar seu padrão de vida e assim promover seu desempenho escolar ou ainda o desafio que existe para desenvolver um processo de ensino-aprendizagem que atenda suas necessidades.

O PNAD 1.02 - Questionário de Educação - 1982 fornece informações sobre a instrução e a ocupação dos ascendentes do chefe e cônjuge do domicílio.⁽¹³⁾ Em princípio, estas informações permitiriam analisar a melhoria no nível de escolarização entre gerações e ainda quais grupos sócio-culturais se beneficiaram mais da expansão da oferta de ensino durante as últimas décadas. O registro da ocupação do pai do chefe e cônjuge oferece possibilidades para estudar a mobilidade ocupacional entre gerações.

Finalmente, deve-se mencionar que o arquivo da PNAD-1982 inclui informações sobre 3738 professores atuantes nas redes de ensino público e particular. Informações dos questionários de mão-de-obra e de educação podem ser utilizadas para desenvolver um perfil de salários, nível de instrução, idade e sexo dos professores nos três graus de ensino, as despesas diretas com pessoal docente ou ainda as despesas diretas por aluno.

Por exemplo, a PNAD revela que as despesas diretas com pessoal docente montaram a US\$84 por aluno-ano de instrução no ensino de 1º grau no ano de 1982. No Nordeste, os salários docentes mal chegaram a 54% do valor médio nacional. As diferenças regionais em salários se devem principalmente ao nível de qualificação do corpo docente nas diversas regiões do País.

Mesmo assim, considerados o nível de qualificação, horas trabalhadas, ramo de ensino, série instruída e uma série de outros fatores, os salários do corpo docente de 1º grau eram 1,6 vezes mais altos no Sudeste urbano do que no Nordeste urbano e 2,3 vezes mais altos no Sudeste rural do que no Nordeste rural. Para mostrar a penúria de alguns dos professores rurais do Nordeste, a PNAD-82 informa que o salário médio mensal dos professores nas áreas rurais do Ceará era de apenas US\$13.

5. O SEEC CENSO DE ESTABELECIMENTOS DE ENSINO – 1982

O SEEC/MEC tem suas origens mais remotas no antigo Serviço de Estatística da Educação e Saúde no início da década de 30, quando pertencia

aos quadros do IBGE. Sob a direção de M. A. Teixeira de Freitas, este Serviço desenvolveu uma série de importantes estudos sobre o ensino brasileiro, publicados na Revista Brasileira de Estatística entre 1940 e 1947. Por volta do início da década de 1950, o Serviço foi transferido para os quadros do Ministério de Educação e Saúde e, a partir de então, o IBGE praticamente desistiu de publicar análises do ensino.

A partir de então, o SEEC assumiu responsabilidade para o desenvolvimento de estudos nessa área e hoje é provavelmente a fonte mais citada em atividades educacionais. Em anos recentes, sua maior responsabilidade envolve o preparo, execução e análise do chamado "SEEC Censo de Estabelecimentos de Ensino", um inquérito anual que em princípio cobre todas as escolas de 1º e 2º graus do País, que hoje excedem 200.000. No trabalho de campo deste inquérito gigantesco colaboram vários governos de Estados no Sudeste e Sul e ainda muitas agências do IBGE que participam na distribuição e recolhimento do material de pesquisa.

A unidade de análise investigada pelo SEEC é a escola, cujo corpo administrativo é responsável pelo preenchimento de um de dois questionários elaborados pelo MEC. O primeiro destes questionários se destina a escolas de uma sala só; o segundo para escolas de maior porte.

Os formulários consistem de uma série de "tabelas" que deveriam ser preenchidas contando todos os indivíduos na escola que pertencem a uma determinada categoria. Os dois formulários contêm uma parte em comum que registra o endereço da escola, visa caracterizar seu funcionamento (dias por semana e turnos por dia, especificar fonte e destinação dos recursos aplicados na escola e ainda contabilizar o rendimento escolar no ensino de 1º grau. O questionário para escolas de maior porte investiga bibliotecas, práticas desportivas, faixas de remuneração do pessoal docente e administrativo, dados sobre o rendimento da educação pré-escolar e do 2º grau (veja Anexo 2).

Tendo a escola como base para o levantamento, o SEEC não dispõe de informações próprias sobre a população em idade escolar fora de escola. Para estimar o tamanho dessa população, o SEEC depende do cálculo da diferença entre estimativas do total da população em idade escolar oriundas de outras fontes e o total de matrículas verificado em seus próprios levantamentos. Sem poder colher estas informações simultaneamente torna-se difícil o uso dos dados num estudo da localização e suficiência da escola.

As preocupações encontradas nos questionários do SEEC refletem uma nítida orientação administrativa-burocrática. Em seu conteúdo temático não se encontram informações sobre suas atividades pedagógicas e alguns detalhes sobre as condições físicas dos estabelecimentos. As únicas informações

disponíveis no Censo de Estabelecimentos de Ensino sobre os alunos referem-se as matrículas e o rendimento escolar (transferências, aprovados e reprovados).⁽¹⁴⁾

Como já discutido, o uso destas informações para calcular taxas de transição levou a valores inconsistentes com outras fontes e metodologias. Suspeita-se que o número de repetentes pode ser subestimado nos dados do SEEC pelas implicações que uma alta taxa de repetência têm no conceito do professor ou da escola. Os dados da estatística escolar poderiam ser maquiados na fonte para proteger estes conceitos e evitar um exame de desempenho.

Mesmo sem questionar a lisura da administração escolar responsável pelo preenchimento dos questionários, cada diretoria apresentaria informações sobre a situação interna da escola. Nestas circunstâncias, os alunos que abandonam o curso antes do final do ano letivo seriam declarados como "evadidos" pela escola, ainda quando não abandonem o sistema de ensino. Portanto, a soma das declarações da administração escolar não revela necessariamente o desempenho do sistema de ensino como um todo. A soma total reflete a agregação de visões parciais.

Seria de interesse usar os dados do SEEC para estudar a remuneração do corpo docente. Mas as informações sobre faixas de remuneração aparecem apenas no questionário destinado a escolas de mais de uma sala; o arquivo de dados do censo de estabelecimentos não contém informações sobre o rendimento do corpo docente em escolas de apenas uma sala de aula. Esbarramos no fato que os dois questionários não contêm as mesmas informações.

Mesmo assim, poderia ser interessante relacionar o grau de formação dos professores com o nível de sua remuneração no caso das escolas de maior porte. Mas estas duas informações aparecem em "tabelas" diferentes no formulário. Portanto, não existe a possibilidade de cruzar estas duas informações. As possibilidades de análise são limitadas pelo formato do questionário.

Parece que os dados do SEEC não permitem uma avaliação do desempenho do sistema como um todo, ao mesmo tempo que não permitem estudos ao nível de indivíduos. Limitam-se apenas a caracterização das escolas. Sem nenhuma referência que indica o nível sócio-econômico das escolas, não podemos considerar o desempenho das escolas no seu contexto social. No conjunto, estas são limitações bastante graves.

Do ponto de vista analítico, os dados do SEEC/MEC são úteis para caracterizar o tipo de curso oferecido nas escolas do Brasil. Os dados do Censo de Estabelecimentos de Ensino revela que 54% das escolas de 1º grau oferecem cursos com apenas a primeira série.

Estas escolas são muito pequenas; portanto, elas representam apenas 14% da matrícula de 1º grau. Mas outros 34% dos alunos estudam em escolas que não vão além da terceira série e 41% em escolas com cursos equivalentes ao antigo curso primário. Apenas 11% dos alunos estudam em escolas que avo mais além (da quinta até a oitava série).

Como os dados do SEEC incluem o nome e endereço das escolas, os mesmos podem ser úteis na seleção de uma amostra de escolas para estudos especiais. Neste particular o arquivo de dados serviu para selecionar escolas para um estudo das condições físicas das escolas, da frequência de diretores, professores e alunos e do serviço da merenda escolar.⁽¹⁵⁾ A mesma fonte pode ser útil para selecionar escolas para outros estudos, sobretudo, aqueles que ainda não existem nas áreas da prática pedagógica dos professores e do domínio cognitivo dos alunos.

A pobreza relativa dos dados do SEEC levanta sérias dúvidas quanto a sua utilidade para fins de planejamento nacional. Em princípio, as PNADs podem fornecer informações sobre o rendimento das escolas a nível regional e nacional. A finalidade dos dados do SEEC seria então o fornecimento de informações a nível de escola, próprias para as administrações estaduais e, de fato, constatamos que estas são, as vezes, as únicas informações sistemáticas sobre as escolas disponíveis em vários estados do Nordeste. A utilidade desses dados para finalidades de administração não justifica a existência de uma unidade de estatística ligada ao Ministério de Educação. A tarefa de manter relações das instituições comerciais e dos serviços sociais é geralmente uma responsabilidade entregue ao IBGE, que dispõe de toda a infra-estrutura necessária para este fim. Nestas circunstâncias, parece-nos que a maior parte das atividades do SEEC poderia ser transferida ao IBGE, sem prejuízo do Ministério de Educação. Isto liberaria o SEEC para atividades de maior utilidade do Ministério, sobretudo, a investigação do domínio cognitivo dos alunos que, por sua natureza, requer metodologias e instrumentais especializados que não são da área de competência do IBGE.

ANEXO 1

PNAD 1.02 -QUESTIONARIO DE EDUCAÇÔ – 1982

Tipo de inquérito: amostra a base domiciliar. Unidade de análise: domicílios e pessoas. Abrangência: nacional, excluído o Norte rural. Periodicidade: noção especificada.

1. Identificação do domicílio
 2. Existência de livros Quantidade e tipos, rádio e televisão
 3. Idade na data de referência
 4. Cor
 5. Para chefe e para cônjuge do domicílio
 - 5.1. Instrução do pai
 - 5.2. Instrução da mãe
 - 5.3. Idade em que começou a trabalhar
 - 5.4. Ocupado no primeiro trabalho
 - 5.5. Ramo desse primeiro trabalho
 - 5.6. Posição no primeiro trabalho
 - 5.7. Ocupação do pai no primeiro trabalho do chefe/cônjuge
 6. Para pessoas de 0 a 6 anos
 - 6.1. freqüência a pré-escola e se é paga ou gratuita
 7. Para pessoas de 4 anos ou mais
 - 7.1. Prática da atividade física (vezes por semana ou então porque não pratica)
 - 7.2. Para pessoas de 4 a 18 anos
 - 7.2.1. Número de horas assistindo televisão
 - 7.2.2. Número de horas ajudando nos afazeres domésticos
 8. Para pessoas de 7 anos ou mais
 - 8.1. Para pessoas de 7 a 9 anos
 - 8.1.1. Participação na PEA e horas trabalhadas
 - 8.2. Caso freqüente curso regular
 - 8.2.1. Série e grau do curso freqüentado
 - 8.2.2. Escola pública ou particular (e o valor da mensalidade quando paga)
 - 8.2.3. Existência e valor de outras taxas
 - 8.2.4. Turno freqüentado
 - 8.2.5. Freqüência de aula de educação física (e quantas vezes por semana ou então porque não freqüenta)
 - 8.2.6. Caso freqüente o primeiro grau
 - 8.2.6.1. Hora de início e término das aulas
 - 8.2.6.2. Tempo gasto no percurso casa/escola
 - 8.2.6.3. Principal meio de locomoção utilizado no percurso casa/escola
- 8.2.6.4. Compra ou disponibilidade dos livros didáticos
 - 8.2.6.5. Trabalhos escolares feitos em casa e sua freqüência
 - 8.2.6.6. Caso freqüente escola pública
 - 8.2.6.6.1. Oferta e consumo da merende escolar
 - 8.3. Caso não freqüente curso regular
 - 8.3.1. Série e grau da última série concluída e de concluiu o curso
 - 8.3.2. Se freqüentou escola publica, particular ou publica e particular
 - 8.3.3. Se tiver menos de 36 anos, não concluiu curso superior e não fez o supletivo
 - 8.3.3.1. Motivo porque deixou de freqüentar escola
 - 8.3.3.2. Se pretende voltar a estudar no ano que vem e o tipo de curso que pretende freqüentar
 - 8.3.3.3. Idade em que concluiu a última série ou curso que freqüentou
 - 8.4. Caso não concluiu e 4º série do primeiro grau
 - 8.4.1. Se sabe escrever o nome completo (ou se já soube)
 - 8.4.2. Se sabe escrever um recado (ou se já soube)
 - 8.4.3. Caso nunca freqüentou curso regular
 - 8.4.3.1. Motivo porque nunca freqüentou curso regular
 - 8.5. Calo freqüentou curso regular
 - 8.5.1. Se freqüentava curso regular no ano passado (1981), qual a série e grau freqüentada naquele ano e se em escola publica ou particular
 - 8.5.2. Caso concluiu o 1º grau
 - 8.5.2.1. Se concluiu o curso através do supletivo
 - 8.5.3. Caso freqüenta ou freqüentou o 2º ou 3º graus
 - 8.5.3.1. Idade em que concluiu o 1º grau
 - 8.5.4. Caso concluiu o 2o grau
 - 8.5.4.1. Se concluiu o 2º grau através do supletivo

ANEXO 2

SEEC CENSO DE ESTABELECIMENTOS DE ENSINO – 1982

Tipo de inquérito: censo.

Abrangência: nacional.

Periodicidade: anual.

Unidades de análise: escolas de 1º e 2º graus.

1. Para todas as escolas
 - 1.1. Identificação da escola (código, endereço etc.)
 - 1.2. Dependência administrativa e tipo de mantenedora
 - 1.3. Situação urbana-rural
 - 1.4. Número de dias em funcionamento durante a semana
 - 1.5. Número de turnos por dia
 - 1.6. Fontes do material didático
 - 1.7. Atendimento da merenda escolar
 - 1.8. Fontes e destinação dos recursos aplicados na escola
 - 1.9. Dados do ensino de primeiro grau
 - 1.9.1. Pessoal docente em exercício, segundo grau de formação
 - 1.9.2. Matrícula inicial, segundo série e idade
 - 1.9.3. Rendimento escolar (transferências, aprovados, reprovados, segundo a série)
2. Para escolas de mais de uma sala
 - 2.1. Número de salas de aula utilizadas
 - 2.2. Presença de biblioteca
 - 2.3. Número de pessoas envolvidas em atividades de orientação educacional
 - 2.4. Número de alunos envolvidos em diversas práticas desportivas, por faixa etária e sexo
 - 2.5. Faixas de remuneração, segundo tempo de serviço e habilitação pedagógica
- 2.6. Dados da educação pré-escolar
 - 2.6.1. Pessoal docente em exercício, segundo grau de formação
 - 2.6.2. Pessoal docente em exercício, segundo situação funcional
 - 2.6.3. Número de turmas, segundo turnos
 - 2.6.4. Matrícula inicial e final, segundo idade
- 2.7. Dados do ensino de primeiro grau
 - 2.7.1. Pessoal docente em exercício, segundo situação funcional
 - 2.7.2. Número de turmas, segundo turnos e séries
 - 2.7.3. Concluintes, segundo idade
- 2.8. Dados do ensino de segundo grau
 - 2.8.1. Pessoal docente em exercício, segundo grau de formação
 - 2.8.2. Pessoal docente em exercício, segundo situação funcional
 - 2.8.3. Número de turmas, segundo turnos e séries
 - 2.8.4. Matrícula inicial e final, segundo série e idade
 - 2.8.5. Rendimento escolar (transferências, aprovados, reprovados, segundo a série)
 - 2.8.6. Concluintes, segundo idade
 - 2.8.7. Número e fonte de bolsas de estudo (escolas particulares)

NOTAS

1. Fundação IBGE, *Metodologia da Pesquisa Nacional por Amostra de Domicílios na Década de 70* (Rio de Janeiro: IBGE, 1981), v. 1.
2. U.S. Bureau of the Census, *Atlântida: Un estudio de caso en encuestas de hogares por muestra*, Series ISPO (Versión en español, (Washington, D.C.: Bureau of the Census, 1967), 14 v. 3. Veja-se U.S. Bureau of the Census, op. cit., Ilustración I I-1-1 e II-1-2, pp. 35-37.
4. Fundação IBGE, op. cit., pp. 23-24.
5. Raul Romero de Oliveira, "A Utilização dos Dados das Pesquisas por Amostragem Domiciliar", *Revista Brasileira de Estatísticas*, v. 31, n. 122 (1970), p. 159.
6. Fundação IBGE, op. cit., p. 67.
7. Fundação IBGE, *Pesquisa Nacional por Amostra de Domicílios - 1982: Brasil e Grandes regiões*, v. 6, t. 1 (Rio de Janeiro: IBGE, 1983), p. xxxvii.
8. Fundação IBGE, *Metodologia da Pesquisa Nacional por Amostra de Domicílios na Década de 70*, op. cit., v. 1, p. 106.
9. *Ibid.*
10. *Ibid.*, p. 114.
11. P.R. Fletcher e S. Costa Ribeiro, "O Ensino de Primeiro Grau no Brasil de Hoje", *Em Aberto* (Brasília, ano 6, n. 33 (jan./mar., 1987), p. 3.
12. Veja o item 8.5.1. no Anexo 1. 13. Veja item 5 no Anexo 1. 14. Veja Tabela 2.
15. C. de Moura Castro e P. Fletcher, *A Escola que os Brasileiros Frequentaram em 1985* (Brasília: (PLAN/IPEA, 1986).

COMENTÁRIO

*Divonzir Arthur Gusso **

A produção recorrente de estatísticas incorre num paradoxo terrível: deve produzir um conjunto de dados úteis ao maior número possível de usuários e deve eleger este conjunto de modo a satisfazer critérios de interdependência e de consistência entre os fenômenos e conceitos expressos pelos dados. Em outras palavras, colocasse uma questão um tanto hamletiana: ter e não ter um paradigma como base de geração do plano de dados.

Nosso propósito, que é debater o trabalho de Fletcher & Ribeiro, tem a ver com essa questão. A PNAD, até onde sabemos, não foi prevista para atender a demanda de dados de modelos de fluxos escolares, nem no questionário principal, nem ao nível do suplemento sobre educação.

* IPEA/IPLAN/CEC

Não obstante, vê-se que é possível atender-las, não obstante ninguém haver pensado que o paradigma orientador do seu plano de dados deveria conter essa possibilidade. De certo modo esta afirmação antecipa um dado de apreciação da PNAD: sendo capaz de atender no geral, dispensou-se de prever o atendimento no particular. Mostra-se útil, portanto, mesmo quando dá resultados não esperados.

Entretanto, quantas vezes mais poderá repetir essa performance? E de que dependem esses "by products" da própria formulação do plano de dados ou da perspicácia e engenhosidade dos usuários? Outra perquirição interessante dirige-se aos aspectos operacionais e organizativos da utilização dos resultados da PNAD. Como é possível explorar ao máximo as virtualidades de arquivos de dados como os da PNAD e dos Censos Demográficos? Que recursos de organização seriam exigidos para obter maior rentabilidade dos investimentos feitos para tornar disponíveis estas fontes de informação social?

Se pudermos tomar o pretexto da discussão deste exemplo tão rico de utilização da PNAD em educação - o trabalho de Fletcher & Ribeiro - para também aprofundar a discussão sobre os temas mais abrangentes das relações entre pesquisa e produção de dados por amostra de domicílios, sairemos bem gratificados deste evento.

PARADIGMAS, MODAS E MODICES

Pode-se partir de uma rememoração das origens das PNADs para situar o tema dos paradigmas que orientam sua constituição. Adiantando, pode-se entender, por paradigma, aqueles corpos de conhecimento que, partindo de conjuntos de proposições primárias, permitem derivar proposições verificáveis, com isenção da adoção de protocolos de dedução.

Pois bem, quando se introduziram na América Latina as primeiras experiências de amostragem de domicílios para gerar estatísticas demográficas e sócio-econômicas, seus planos de dados eram formulados segundo as necessidades de informação de modelos de avaliação de disponibilidades e da utilização de recursos humanos, fundados, por sua vez, em paradigmas conhecidos como "manpower approach", as vezes hibridados com os de "capital humano". Ou, ainda, em alguma variante de teorias neoclássicas do emprego.

Subjaziam a estas opções conceituais e metodológicas preocupações com problemas práticos de política governamental, muito marcados pelas condições sociais de meados dos anos sessenta: crescimento acelerado e conseqüentes desajustes nos níveis de emprego e de estabilidade econômica;

mobilidade mais intensa de trabalhadores e de capitais; descompassos na oferta de educação e qualificação de mão-de-obra diante de variações nas demandas setoriais e assim por diante. Nos países periféricos, essa problemática era aguçada e peculiaridade pelas maciças migrações rural-urbanas e intersetoriais, elevadas taxas de crescimento populacional e acentuada transição sociocultural, crescentemente desigual distribuição de renda e inquietantes níveis de pobreza.

Evidentemente, o plano de dados teria que incluir a expressão dos fenômenos observados segundo os conceitos e protocolos de mensuração criados segundo esses paradigmas. E objetivariam alimentar pesquisas por eles orientadas.

Logo, porém, o poder explicativo desses paradigmas e sua correspondente maneira de problematizar a realidade começariam a ser contraditados por novos paradigmas ou teorias parciais emergentes. Como, por exemplo, a explicação dos desajustes de emprego e renda pela dualidade das estruturas produtivas e pela segmentação dos mercados de trabalho. Do mesmo modo os fenômenos associados à pobreza passam a receber novos enfoques, com a superação, no campo da economia e da sociologia, das teorias dominantes a respeito da modernização, da mobilidade social e da integração aos mercados de trabalho e da formação e distribuição dos rendimentos familiares.

No tocante a educação, como campo disciplinar derivado, assiste-se a uma passagem das abordagens economistas - onde o capital humano é o paradigma hegemônico - para explicações fundadas nas funções sinalizadoras e socializadoras (ainda dentro das matrizes teóricas estrutural-funcionalistas) nas peculiaridades dos padrões de exclusão social inerentes ao desenvolvimento dependente e nas determinações reprodutivistas das superestruturas capitalistas (já em paradigmas divergentes dos anteriores). Todavia, com a própria superação da crença no papel explicativo das variáveis educacionais sobre os fenômenos de distribuição da renda e da modalidade social, parece declinar a importância "em si" das características educacionais da população na constituição do plano de dados das pesquisas domiciliares.

Talvez por isso se observe que quase nada muda em sua configuração conceitual e no espaço que ocupa nos instrumentos de coleta.

Não obstante, é pertinente atribuir este fato também a outra circunstância. Uma fonte de dados sociais e sua constituição evoluem na medida da avaliação objetiva e concreta de sua utilidade, que só se pode dar a partir da prática da investigação social nelas fundada; que poderia indicar insuficiências, impropriedades e lacunas em sua capacidade de suprir os modelos de análise que se vão construindo segundo o apuramento dos processos de problematização.

Pelo menos no caso brasileiro, é possível evidenciar que os principais usuários dos arquivos PNADs foram os pesquisadores das áreas temáticas de emprego, salários e distribuição de renda, desenvolvimento regional e urbano, migrações internas e mercados de trabalho. Igualmente importantes foram as demandas surgidas no início dos anos setenta, quando se disseminou a moda dos "indicadores sociais" como um dos possíveis instrumentos de desenho de políticas de desenvolvimento social: percebe-se que esta é uma nova fonte de estímulo ao aperfeiçoamento das fontes de dados sociais - como o demonstram as mudanças de configuração do plano de dados da PNAD a partir de 1976 e o lançamento da ENDEF.

Embora sejam poucos os registros formais e sistemáticos da produção de pesquisas em educação para essa época a possível afirmar com alguma segurança ser escassa, para não dizer rara, a utilização destes dados na pesquisa educacional. Os usuários estiveram, por certo, concentrados na produção de tipo mais pragmático de "diagnósticos", "estudos para o planejamento" e fundamentação quantitativa de políticas e de projetos de investimentos.

É preciso notar que, no campo da pesquisa educacional, antes de ciclos ou hegemonias de paradigmas, parece haver predominado, em nosso País, de um lado, uma forte dispersão temática e teórica e, de outro, uma não surpreendente adesão diletante aos modismos "teóricos" em voga nos meios acadêmicos metropolitanos. Em verdade, a aceitação ou rejeição de novos ou velhos paradigmas que se confrontavam nos países centrais, se dava menos por esforço de pesquisa e verificação do que pela discussão periférica e pela crítica externa dos fundamentos teóricos (e algumas vezes até das bases epistemológicas) dos paradigmas dominantes em cada corrente ideológica ou acadêmica.

Caberia lembrar, de passagem pelo menos, que vigorou por muito tempo (e ainda persiste) uma certa confusão entre pesquisa fundada em verificação empírica (com perdão do virtual pleonasma) e pesquisa orientada por concepções positivas. De que decorre um certo "sentimento de culpa", entre cientistas sociais e pesquisadores de educação, quando, como ocorre freqüentemente, se vêem compelidos a preparar uma tese ou a realizar uma investigação baseada em "levantamento de campo" ou lastreada na análise de estatísticas secundárias. E certamente não se deve olvidar o fato de que, em nossas academias, é escasso o treinamento em métodos e técnicas de pesquisa social, que permitiria "transar" sem culpa (e, pois, pressupondo competência) procedimentos mais sofisticados de investigação, que implicam o uso de técnicas quantitativas sem abdicação de posturas epistemológicas não-positivistas ou mesmo anti-positivistas.

Sem dúvida este não é um quadro estático, imutável. Apesar da superficialidade das modices, os problemas educacionais do País tem uma incontrastável dinâmica e colocam exigibilidades novas à produção de conhecimentos a seu respeito. Quando menos, à medida em que vão fracassando muitas políticas públicas e movimentos político-ideológicos que tentam influir sobre elas e vão emergindo novos conflitos e demandas políticas diferenciadas.

USOS POTENCIAIS E POTENCIAL DE USO

Talvez um exemplo provocante de como uma configuração de dados não prevista para um determinado fim acaba por atender-se seja justamente o da pesquisa de Fletcher & Ribeiro sobre a extensão da escolaridade de diferentes grupos na sociedade brasileira e sobre a eficiência interna dos processos escolares.

Tradicionalmente estes dois problemas tem sido estudados separadamente e com base em fontes diversas de dados: o primeiro sobre as medidas de danos de estudo e curso completo da população captados nos censos demográficos e o segundo sobre os registros de eventos escolares captados em censos de estabelecimentos processados pelo SEEC/MEC e, episodicamente, em levantamentos "ad hoc" realizados para projetos específicos de pesquisa. E, quase sempre. Sob quadros conceituais e técnico-metodológicos pouco explícitos e precariamente apurados.

Em parte, a inovação do tratamento uniforme e articulado foi possibilitada pelo Suplemento da PNAD de 1982. Mas teria sido possível fazê-lo com um aperfeiçoamento pouco complicado do formulário básico, de vez que a maior parte dos quesitos do suplemento nada têm a ver com o quadro de variáveis incorporado ao modelo de Fletcher & Ribeiro.

Predomina, neste, a engenhosidade em combinar dados e calcular medidas derivadas de variáveis mais ricamente concebidas no contexto de um modelo mais integrado e totalizador. E uma certa ousadia ao procurar superar uma velha e sólida crença em argumentos do tipo "o problema fundamental da escola brasileira é a evasão e a repetência", com a demonstração objetiva de que a evasão é bem menor do que usualmente se admite, mesmo na 1ª série do ensino fundamental. Demonstra, também, que as sucessivas reprovações e a acumulação de repetências, estas sim, levam ao abandono precoce da escolaridade. E consegue medir os momentos em que, tendencialmente, estes eventos ocorrem.

Com efeito, a incidência e a dimensão destes fenômenos é característica das imagens de fluxos escolares com que se procura caracterizar a problemática da escolarização insuficiente do brasileiro. Imagem, aliás, nitidamente "ideológica" na medida em que se coloca como a explicação social que orienta a prática das políticas de educação.

Longe porém de constituir-se no problema, estas categorias são manifestações ou exteriorizações igualmente identificadas e classificadas de fenômenos diferentes que constituem, estes sim, os verdadeiros problemas a serem enfrentados: a apropriação social desigual das oportunidades escolares que o sistema educativo brasileiro é capaz de oferecer.

Isto se evidencia de modo claro nos "findings" de Fletcher e de Fletcher & Ribeiro, ao mostrarem haver uma certa invariância das medidas de reprovação quando varia a localização e o nível de renda ou de situação social. Nos espaços sócio-educativos destinados aos estratos sociais "altos" a reprovação pode ser elevada mas representa maiores exigibilidades para promoção; naqueles destinados aos estratos "baixos" a reprovação se daria com menores exigibilidades de "conteúdo" e se relacionariam mais fortemente com os efeitos de exclusão.

Mesmo limitando comentários substantivos a este tópico, é possível afirmar que a pesquisa de Fleete & Ribeiro constitui um bom exemplar de estudo descritivo e analítico dos efeitos da escolaridade socialmente desigual que se pratica em nosso País. Revela, objetivamente e sem preconceitos, como se manifestam as desigualdades de acesso às oportunidades de escolarização e como elas se projetam na trajetória escolar discriminatória de alunos procedentes de diferentes estratos sociais e diferentes espaços sócio-econômico. E contribuem, desde logo, para estudos dedicados a formulação de políticas públicas e de planos de governo, na medida em que a estrutura do modelo analítico permita, potencialmente ao menos, expansões para modelos de simulação e de prospeção.

Não obstante, a estrutura do modelo e as mensurações obtidas não se encontram ainda cabalmente exploradas. E podem exigir algum apuramento para abrir novas sendas de explicação dos fenômenos de escolaridade desigual. Importante a observar, é que isto ainda é possível nos quadros mesmos dos arquivos da PNAD.

Conquanto apresente boas medidas dos "outputs" e das relações internas de processamento escolar, o modelo oferece poucas informações a respeito dos nexos entre essas medidas e o encontro social e características mais fortemente identificadoras dos grupos sociais que a escola atende. Sem dúvida é um avanço haver medido e parametrado a classificação "social" dos educandos mediante um indicador de estrato social; mas exatamente aí reside uma fragilidade: renda, ocupação dos membros do domicílio e características

de dotações domésticas - como menciona Lara Resende - precisam ser cuidadosamente tratadas para refletir adequadamente as posições e distancias sociais.

Um estudo recente de Behrmann & Birdssal abre um caminho para considerar esse entorno e características de "inputs" escolares como "proxis" da qualidade da escolarização que alteram substancialmente os resultados de análises da quantidade de escolarização que um sistema oferece. Um esforço nessa direção não apresentaria grandes dificuldades como já se demonstrou em um processamento de dados da PNAD, também realizado por Fletcher a pedido da Coordenação de Educação e Cultura do (PLAN/IPEA, sobre características das pessoas ocupadas como professores no ensino de 1º e 2º grau (salários, tempo de trabalho semanal, status sócio-econômico, nível de instrução e tipo de escola)

Outros estudos poderiam ser derivados daí, como os que relacionassem os eventos de freqüência, trajetória e "outputs" com características da entrada no mercado de trabalho e da geração dos rendimentos familiares. Ou as associações possíveis entre aqueles eventos e a mobilidade espacial e ocupacional dos membros do grupo domiciliar.

Avanços mais substanciais, no entanto, dependeriam de interações e articulações mais amplas entre pesquisas baseadas em dados de amostras domiciliares com outras baseadas em amostras de unidades prestadoras de serviços sociais (escolas, postos de saúde etc.) que mostrassem, num mesmo espaço social (possibilitando a associação de entornos), os eventos educacionais e ambientais.

Esta observação implica em afirmar que não se deve esperar de um numero muito grande e diversificado de eventos e variáveis numa só operação de registro e coleta de dados. A PNAD demonstrou ser portadora de um fértil e seguro conjunto de métodos e técnicas de obtenção de dados sociais e econômicos; mas também demonstrou que tem seus limites e insuficiências, tanto no plano técnico (como o de capacidade da amostra) como no conceitual (o que pode e não pode ser expresso num quesito ou grupo de quesitos).

Um exame mais acurado dos planos de dados e dos temas de investigação que se vão demonstrando relevantes (e necessários) desvendaria, por certo, potencialidades de uso notavelmente estimulantes. Mas aí é preciso indagar qual o potencial de uso que iremos encontrar nas organizações de pesquisa educacional e de produção de informações para os sistemas de tomada de decisão no âmbito publico e privado, ou seja, no complexo universo dos usuários.

Cumpre ver que os critérios e processos de acesso dos pesquisadores aos arquivos de dados sociais - seja os da PNAD ou do SEEC/MG não

facilitam de modo algum uma utilização mais ampla dessas fontes.

Tampouco os planos de ensino e de treinamento para pesquisa estimulam a exploração de técnicas quantitativas e o recurso a métodos mais *rigorosos* de observação empírica; sequer ensinam o "caminho das pedras" para fazer uso mínimo dos arquivos existentes.

À GUIA DE CONCLUSÕES

Após reafirmar que o modelo Fletcher & Ribeiro é estimulante e merece ser melhor testado, avaliado e expandido, cabe sugerir que, na mesma senda, se pudesse conceber alguns avanços na própria fonte de dados. O modo como foi planejado o Suplemento de 1982 a tardia utilização que se deu a seus dados, indicam, por exemplo, como não se deve agir. A título de participação ampliada dos usuários, abriu-se campo para um caleidoscópio de demandas sem compromisso de posterior utilização. Indagou-se de pré-escola, educação física, merenda, transporte escolar; e estudou-se apenas fluxos escolares que não tiveram tratamento mais específico no plano de dados.

Se fosse aceita a idéia de um programa de suplementos, ao longo dos anos noventa, para vários temas setoriais (já que esta é a preferência visível), conviria organizar um grupo de trabalho, formado por instituições usuárias selecionadas, que, a partir de um programa de pesquisas básico (programa *umbrella*), gerariam o plano de dados de cada suplemento. Os participantes desse programa comprometer-se-iam formalmente a concretizar o conjunto de análises previsto num prazo determinado, concorrendo para a cobertura dos custos de processamento.

Uma expansão dessa idéia, seria considerar a possibilidade de, em cada suplemento, serem investigadas tanto as variáveis típicas do "setor" como algumas de outros setores que a ela se relacionam, permitindo acumular, na década, um conjunto razoável de relações intersetoriais entre características da população e dos grupos sociais que a compõem.

Ademais disso, em lugar da PNAD se poderia ter um "sistema" ou um "network" de pesquisas por amostras de domicílios e de unidades de serviços sociais, a cargo do IBGE e dos sistemas setoriais de estatísticas, que permitisse redistribuir as várias demandas por dados entre levantamentos distintos de menor porte e complexidade e, pois, com menores custos de coleta e processamento.

Em contrapartida, faz-se necessário um esforço conjunto de produtores de estatísticas e responsáveis pelos centros de treinamento de pesquisadores acadêmicos e de agências governamentais para disseminar e incrementar

competências para o uso de técnicas quantitativas e exploração dos arquivos e bancos de dados disponíveis e para flexibilizar e tornar mais eficientes os canais de acesso as fontes. E naturalmente, para proporcionar um sistema mais eficaz de participação do usuário no planejamento das fontes que, afinal, só tem razão de existir se houver de fato este usuário.

REFERENCIAS BIBLIOGRÁFICAS

BEHRMANN, Jere & BIRSA, Nancy. *The Quality of schooling. Quantity alone is misleading*; Washington, World Bank (World Bank Rarprint Series nº 311), 1984

DEMO, Pedro. *Avaliação Qualitativa*; São Paulo, Cortez, Autores Associados, 1987

FLETCHER, Philip. *Modelo matemático da trajetória escolar, da repetência e do desempenho do sistema de ensino de 1º grau no Brasil*; Brasília, CNRH/IPLAN/IPEA (Projeto PNUD/OIT), 1985

--- & CASTRO, Cláudio M. *Os mitos, as estratégias e as prioridades para o ensino de 1.º grau*; Brasília, CNRH/IPLAN/IPEA, 1985

-- & RIBEIRO, Sérgio C. *A educação na estatística nacional*; Rio de Janeiro, Paper apresentado ao Seminário de Avaliação das PNADs da Década de 80, realizado pela ABEP em Nova Friburgo, 13-15 de junho de 1988.