

Investimento em educação no Brasil: uma réplica

CLÁUDIO DE MOURA CASTRO

Na última edição desta Revista publicamos um artigo comentando três estudos de economia da educação realizados no Brasil.¹ Langoni, discordando do que dissemos sobre a sua tese,² defende-se no comentário incluído nesse número, fazendo, ademais, algumas críticas à nossa própria tese,³ que fôra discutida no artigo.

Como ambas as teses ainda estão em processo de publicação, o leitor não tem por onde se louvar senão no que saiu na Revista. Não obstante, o debate é público e direto, com tantos juizes quantos leitores se dispuserem a acompanhar êste tipo de discussão.

Em divergências técnicas, a rigor, alguém está errado; no entanto, podemos vislumbrar aqui algo que tem mais o sabor de controvérsia doutrinária e que, conseqüentemente, depende de opinião e julgamento. Trataremos inicialmente desta questão.

Diante de um problema em que nossos conhecimentos são particularmente escassos e rudimentares, existem dois tipos de enfoque. Podemos fazer uma análise incompleta do todo, ou conhecer bem uma pequena parte, que não sabemos se representativa.

Ao contrário do que Langoni sugere, num estudo de caso o interesse não é o caso em si (Itabirito) mas sim aquilo que êle nos sugere a respeito do todo Brasil.⁴

¹ Cláudio M. Castro, "Investimento em Educação no Brasil. Comparação de três estudos", *Pesquisa e Planejamento*, junho 1971, n.º 1.

² Carlos Langoni, *A Study in Economic Growth: The Brazilian Case*, Tese de Ph.D. para a Universidade de Chicago, 1970.

³ Cláudio M. Castro, "Investment in Education in Brazil: A Study of two Industrial Communities", Tese de Ph.D. para a Universidade de Vanderbilt, 1970.

⁴ Aliás, foi desta forma que foi entendido nosso estudo pelo Professor T. Schultz da Universidade de Chicago.

Os estudos de caso têm uma tradição muito mais longa em ciências sociais. Por isso mesmo, é óbvio que não requerem justificação; já os estudos agregados, apenas iniciados nas últimas décadas, ainda encontram resistências na profissão (v.g. o tom defensivo do artigo citado por Langoni: "Is aggregation necessarily bad?"). Num estudo de caso, ao invés de apresentar um teste de representabilidade, o pesquisador deixa as inferências relativas ao todo por conta da capacidade de julgamento do leitor, daquilo que Pascal chamou de *esprit de finesse*.

Não parece que esta seja uma controvérsia em que posições *a priori* sejam justificáveis. Em cada caso concreto, o pesquisador terá que se definir por um certo nível de agregação e é aí que emergem as divergências. Langoni opta por um estudo a um nível máximo de agregação (Brasil); de nossa parte, optamos pelo mínimo de agregação.⁵

Havíamos julgado injustificável um alto grau de agregação, tendo em vista que: i) não há evidência empírica de que os pressupostos necessários para o cálculo das taxas de retorno sejam aceitavelmente realistas; ii) os dados existentes são de qualidade desconhecida; iii) para algumas variáveis cruciais os dados simplesmente inexistem a nível agregado (inteligência, variáveis de *status* etc.); iv) apenas com grande risco pode-se conduzir uma análise estritamente econômica de questões educacionais sem um conhecimento aceitável dos parâmetros institucionais e sem perder de vista as dimensões não econômicas da educação; tais exigências aumentariam extraordinariamente os encargos de um pesquisador que opte por uma análise em nível nacional.

⁵ Por outro lado, a escolha de Itabirito nos parece inteiramente justificável sem argumentos subjetivos. Esta é uma cidade *estritamente* industrial (sete pessoas no setor secundário para cada pessoa no primário), que se industrializou no princípio do século, tendo, portanto, superado os desequilíbrios passageiros inerentes ao processo de transformação estrutural. Localizando-se na região aurífera de Minas Gerais, está entre os mais antigos núcleos brasileiros de classe média e trabalhadores livres. Em termos de desempenho educacional, Itabirito apresenta uma performance superlativa (dentre 29 municípios de tamanho equivalente, Itabirito tem mais alunos de nível secundário do que 25 deles). Por outro lado, o município apresenta uma estrutura "pluralista", com mecanismos de *decision-making* descentralizados.

Examinaremos a seguir alguns pontos específicos:

1. Qualidade dos dados

Grande parte da discussão diz respeito à amostra e à qualidade dos dados.

- a) A representatividade da amostra e a fidedignidade das respostas.

Se estimarmos a idade de uma mulher baseados em uma pergunta feita a ela, possivelmente estaremos cometendo um erro cuja magnitude, ademais, não pode ser facilmente avaliada. No entanto, como estamos lidando com o universo, o *erro de amostragem é zero*.

Langoni concentra suas observações na questão dos erros de amostragem, mostrando a contento que na PNAD estes usualmente não passam de um a três por cento. A questão, entretanto, é muito mais simples: sabe-se que existe uma grande relutância em prestar informações a respeito de rendimentos; apenas em pesquisas feitas com este objetivo específico justifica-se tomar tôdas as precauções visando a minimizar o falseamento das respostas.⁶ É, portanto, ocioso saber se o erro de amostragem é 0,7 ou 0,9 por cento *quando sabemos que erros de interpretação, omissão e má fé são muitas vêzes superiores a este montante*. Por outro lado, não conhecemos a magnitude numérica do "bias", que deve diferir em cada nível de educação.

⁶ O autor representou o IPEA na Comissão Censitária Nacional e em diálogos com a equipe do IBGE pôde constatar o pouco crédito que merecem as estatísticas de renda pessoal (Censo ou PNAD).

O Censo e mesmo a PNAD têm um interesse subalterno nos dados de renda — *apenas uma pergunta é incluída no questionário* — estes são inquéritos de utilidade múltipla.

Em Belo Horizonte e Itabirito, por outro lado, a estimativa de renda foi a ênfase do *survey*. Os enumeradores foram alertados para a importância máxima deste tópico e o questionário desmembra a pergunta em várias categorias (renda de "biscates", horas extras, etc.). Para o controle das respostas o questionário inclui nove escalas de *status* sócio-econômico. Todos os esforços foram enviados para evitar o falseamento das respostas e a experiência de campo, de fato, mostrou que mesmo em condições favoráveis algumas categorias de indivíduos dificilmente fornecem respostas seguras.

b) A representatividade da amostra e a estratificação.

Para que, a partir de uma amostra estratificada por nível educacional, possamos estimar, por exemplo, a média de renda da população, é estritamente necessário que o tamanho de cada subamostra (correspondente a um estrato) seja proporcional à participação daquele estrato na população. Langoni discute esta questão em sua crítica.

No entanto, a comparação das médias de cada estrato *não requer proporcionalidade na estratificação*.⁷ Até pelo contrário, se gerarmos amostras maiores para aqueles grupos de maior variância estaremos diminuindo o erro de amostragem da diferença de médias de rendimentos.⁸

Enquanto houver aleatoriedade dentro de cada estrato, e esta condição foi satisfeita em nosso *survey*, perfis de renda estimados por mínimos quadrados não terão senão erro de amostragem.⁹

⁷ Desde que haja aleatoriedade na amostragem dentro do grupo, a média, de renda, digamos, dos analfabetos, será um estimador justo da média da população de analfabetos; o erro estará acima da média estimada com a mesma probabilidade de estar abaixo.

⁸ Em nosso *survey* tentamos fazer a amostra proporcional à variância da renda que é o parâmetro que nos interessa. Tivemos que usar métodos aproximados simplesmente porque esta variância não é conhecida (nem de nós nem do IBGE). Qualquer pesquisa no Brasil que tentasse minimizar os erros nas estimativas de diferenciais de renda teria que usar os métodos aproximados que nós utilizamos.

⁹ Tanto em Belo Horizonte quanto em Itabirito quase todas as amostras foram coletadas aleatoriamente, o que garante que não há viés nos estimadores de renda de cada grupo. A amostra de analfabetos em Itabirito não é aleatória apenas em um sentido trivial. De fato, como Langoni nos cita, "we had to bypass the random sampling procedure"; restou mostrar, como está explicado na mesma página (p. 381), que o objetivo passou então a ser o de entrevistar toda a população de analfabetos. Em uma cidade pequena, com entrevistadores locais isto pode ser praticamente conseguido.

Os testes de significância e os erros padrão de estimativa das regressões de renda estão nas tabelas das páginas 164 e 218-219, mostrando que são todos os ajustamentos significativos pelo menos a 5%; exceto os analfabetos, que por não terem praticamente aumentos de rendimentos durante sua vida, têm necessariamente linhas de regressão que não são significativamente diferentes de uma linha horizontal.

2. As estimativas de custos diretos

Langoni afirma em sua tese que “the estimation of current costs did not present any special problem”.¹⁰ Parece-nos, entretanto, que a facilidade encontrada decorre de uma aceitação de estatísticas particularmente precárias. No rodapé da Tabela 86 constante do Apêndice (p. 209) de sua tese, podemos verificar que os custos baseiam-se na média dos dispêndios de 1968/69/70 e que os dados foram obtidos de *projeções* do Plano Decenal de Desenvolvimento Econômico e Social publicado em março de 1967.¹¹ Consultando o Plano, observamos que as projeções louvam-se nos dispêndios até 1965, e, nas matrículas até 1962 para a educação primária e 1965 para outros níveis. Os autores do Plano deixaram claro o caráter rudimentar das estimativas.¹²

Concordamos inteiramente com Langoni quando êste afirma que os dados de custos de capital são precários.¹³ Em Minas Gerais observamos que escolas são construídas por cinco organismos públicos diferentes, sem que exista uma contabilidade unificada ou possibilidade de desemaranhar os poucos dados existentes. Por esta razão, levantamos os custos de capital, diretamente, em nível de dispêndio, no canteiro de obras de escolas padronizadas, construídas em diversos municípios, o que já é uma aproximação melhor.

3. Perfis de rendimentos

Tanto nos dados do Censo como nos da PNAD, as estimativas de renda pessoal são dadas por intervalos de renda e de idade, e não, em séries contínuas. Aparentemente, isto não se constituiria

¹⁰ *Op. cit.*, p. 122.

¹¹ A referência não traz o ano de publicação do Plano.

¹² *Plano Decenal. Op. cit.*, pp. 147-149.

¹³ Exatamente por isto o esforço de Levy constitui uma contribuição importante, ainda que não pretenda ser a palavra final. Sua avaliação dos investimentos públicos em educação no Estado de São Paulo possivelmente é a melhor estimativa disponível do capital imobilizado em prédios e instalações. Justamente por ser um campo inexplorado, o próprio Levy reconhece as óbvias limitações dos resultados (como indica Langoni em seu “Comentário”).

em uma limitação séria. No entanto, para altas taxas de retorno, digamos 28%, os anos iniciais têm um valor presente altíssimo comparado com os outros. Os intervalos iniciais de idade são de 10 a 14 e de 15 a 19 anos. Sabemos que menores, por lei, recebem 1/2 e 2/3 do salário-mínimo aos 16 e 17 anos, respectivamente, e que este salário configura um mercado de compradores. Os intervalos de idade obliteram completamente grandes diferenças, obrigando Langoni a reconstruir o trecho mais importante do perfil de rendimento em base algo arbitrária, tal como mencionamos no artigo.¹⁴

4. O sistema educacional brasileiro

O sistema educacional brasileiro foi insuficientemente estudado e as fontes de referências são dispersas e de acesso duvidoso.¹⁵ Assim, torna-se difícil fazer justiça à complexidade institucional que certamente condiciona muitas conclusões e inferências.

Apesar de seguidamente fazer sugestões de política educacional Langoni demonstra pouca familiaridade com a realidade brasileira quando menciona a possibilidade de aumentar as matrículas "by more intensive utilization of the stock of physical and human capital".¹⁶ Ora, sabemos que, exceto em nível universitário, onde a situação não pode ser discutida em termos tão esquemáticos, aumentos consideráveis de matrículas provavelmente só poderão ser conseguidos mediante a redução no número de horas de aulas diárias,¹⁷ ou o sacrifício da qualidade da educação que, de resto, já é deficiente.

¹⁴ Em nossas estimativas para graduado em ginásio em Itabirito, várias hipóteses de renda não recebida, todas elas plausíveis, fazem flutuar a taxa de retorno entre 12 e 20% (*op. cit.* p. 235), nota-se que em termos de investimentos em educação estas são taxas modestas. Taxas de retorno mais elevadas, como os 28% encontrados por Langoni, sofreriam flutuações ainda mais amplas.

¹⁵ É de se notar, entretanto, que a bibliografia da tese de Langoni não inclui uma só publicação sobre educação ou mesmo sobre o Brasil.

¹⁶ Página 124.

¹⁷ Nossas escolas primárias operam com dois e três turnos e oferecem em geral de três a quatro horas de aula por dia. As escolas secundárias não estão em situação muito diferente.

Adiante, Langoni discutindo a estrutura de custos da educação afirma que os custos por aluno “rise with the level of education, reflecting mainly the large differences in enrollment as we move across the education spectrum”.¹⁸ Em outras palavras, êle estaria afirmando que universidade custa mais do que primário principalmente porque se opera em uma escala menor, isto é, existem menos universitários do que alunos no primário. Isto nos levaria a afirmar que uma expansão relativa da matrícula nas universidades faria os custos se aproximarem dos do primário. É sabido que os custos de mão-de-obra absorvem cêrca de dois têrços dos custos em educação e que uma professora primária ganha entre cinco e dez vêzes menos do que um professor universitário. Perguntaríamos a Langoni, como será possível diminuir esta diferença de custos?

5. O efeito da educação e as variáveis não controladas

Segundo Langoni, a um “único tipo de viés que poderia ser atribuído *a priori* às estimativas de taxas de retôrno dos investimentos em educação com o uso de dados agregados, não foi discutido por Castro”.¹⁹ Êste viés está longe de ser o único; outros mais importantes já havíamos discutido em nosso artigo. Na medida em que educação e idade são as variáveis independentes disponíveis no Brasil para estudos agregados desta natureza, sabemos que o nível educacional atingido por um indivíduo está fortemente correlacionado com várias outras variáveis, importantes na determinação da renda, falsificando a medida da verdadeira associação entre renda e educação. Na sociedade norte-americana que é muito mais homogênea, apresentando menores desigualdades sociais e culturais, Denison estimou que o efeito destas variáveis reduziria para dois têrços os rendimentos imputáveis à educação.²⁰ Sendo esta diversidade cul-

¹⁸ *Op. cit.* p. 125.

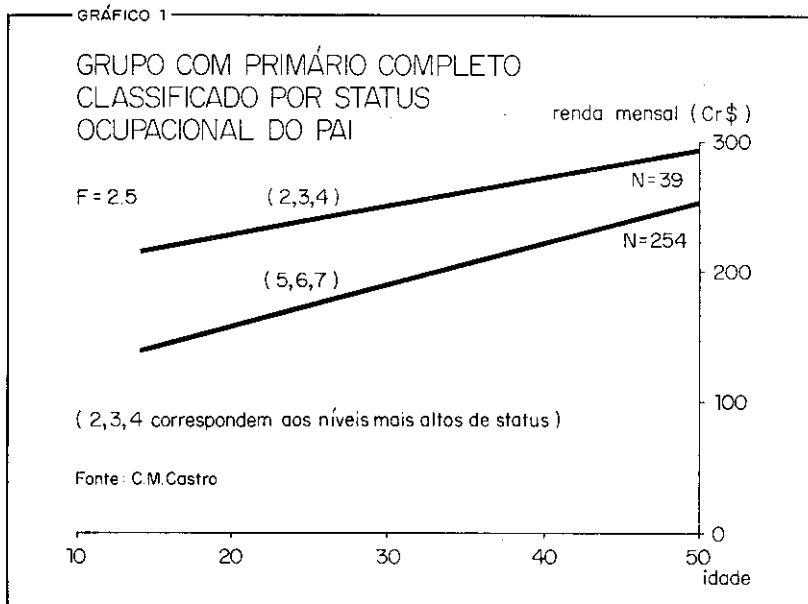
¹⁹ “Comentário ...” *op. cit.* p. 12 (?).

²⁰ E. Denison, *The Sources of Growth in the United States and the alternatives before us*, (New York: Committee for Economic Development, 1962).

tural maior no Brasil,²¹ as taxas de retorno à educação ficam mais superestimadas, e, como conhecemos menos nossa realidade, mais incertas.

Em Itabirito observamos que inteligência e nove variáveis de *status* sócio-econômico não se distribuem aleatoriamente por nível de educação. Para o mesmo nível de educação da amostra, indivíduos com pais de maior *status* obtêm níveis significativamente mais elevados de renda, o mesmo se dando com inteligência e diversas outras variáveis.

O Gráfico 1 exemplifica tal tipo de resultado, que se repete sistematicamente em todos os grupos de Itabirito e Belo Horizonte. Isto



²¹ Que tal diversidade é muito mais ampla no Brasil pode ser facilmente verificado. Podemos, por exemplo, comparar as diferenças de médias de inteligência para classes sociais "alta", "média" e "baixa" nos Estados Unidos (B. Berelson e G. Steiner, *Human Behaviour* (N. York: Harcourt, Brace and World, 1964 p. 60) com o Brasil (P. Wcil, *Pesquisa sobre o nível mental da população brasileira*, SENAC, Deptº. Nacional, 1959; e também C. M. Castro, *Investment*, *op. cit.* Cap. VIII. Em nossa pesquisa mostramos que a diferença da média de inteligência entre crianças de sete anos filhas de analfabetos e universitários é da ordem de dois e meio desvios padrão.

basta para mostrar que ainda que os dados de renda sejam precisos, a relação entre educação e renda é significativamente falseada por uma análise estatística que inevitavelmente ignora certas variáveis.²²

Em Itabirito, os diferenciais de renda superestimam o efeito líquido da educação. Como discutimos na “Comparação...” tampouco resistiu ao teste empírico a hipótese singela de que investimentos substanciais em educação gerariam um aumento notável no crescimento econômico ou alterariam os padrões de distribuição de renda.²³ Em Itabirito a realidade é muito densa e complexa; as análises mais profundas revelam erros sérios nas formulações simples iniciais. Deixamos ao leitor julgar se no Brasil seria diferente. A guisa de resposta à pergunta de Griliches e Grunfeld “Is aggregation necessarily bad?” diríamos, com K. Pearson: “Antes nos satisfazermos com uma fração da solução certa do que nos iludirmos com o todo de uma solução errada”.²⁴

²² Como advertimos em nosso artigo, não estamos afirmando que os retornos ao investimento em educação não sejam importantes, meramente dizemos que a análise inocente da relação renda-educação deixa para trás coisas que não podem ser ignoradas em formulação de política educacional.

²³ Discutindo a situação da Inglaterra e dos Estados Unidos, John Vaizey tampouco pode vislumbrar relações tão nítidas “The correlation between education and lifetime earnings, while it certainly exists, does not of itself prove that education causes the earnings, and there are strong theoretical reasons for holding that at the very least, the point is unproven... Educations’s effects are more subtle than has been supposed by the proponents of the view that it directly affects income distribution, and to suggest that education tends to re-inforce, rather than to diminish existing social inequalities” Manuela F. Leite. P. Lynch, J. Vaizey *et alia* (*The Economics of Educational Costing*, Lisboa, Centro de Economia e Finanças, Instituto Gulbenkion de Ciência, 1969, p. 59 e 73).

²⁴ K. Pearson, *The Grammar of Science*, (London: J. M. Dent & Sons, 1937) p. 36.

