

Métodos Quantitativos na Educação Comparada e em Outros Cursos: são válidos?

Steven J. Klees¹

¹University of Maryland (UMD), College Park/MD – Estados Unidos da América

RESUMO – Métodos Quantitativos na Educação Comparada e em Outros Cursos: são válidos?¹ A comparação é a essência da ciência e o campo da educação comparada e internacional, como muitas ciências sociais, encontra-se dominado por abordagens metodológicas quantitativas. Este artigo levanta questões fundamentais a respeito da utilidade da análise de regressão para inferências causais. Examina três obras extensas de análise de regressão aplicada referentes a políticas educacionais. O artigo conclui que o alcance ou até mesmo a proximidade das condições necessárias para que a análise de regressão produza inferências causais válidas está mais distante do que nunca, de modo que estas inferências nunca são válidas. As metodologias de pesquisa alternativas são então brevemente discutidas. Palavras-chave: **Educação Comparada. Métodos de Pesquisa. Análise de Regressão. Inferência Causal.**

ABSTRACT – Quantitative Methods in Comparative Education and Other Disciplines: are they valid? Comparison is the essence of science and the field of comparative and international education, like many of the social sciences, has been dominated by quantitative methodological approaches. This paper raises fundamental questions about the utility of regression analysis for causal inference. It examines three extensive literatures of applied regression analysis concerned with education policies. The paper concludes that the conditions necessary for regression analysis to yield valid causal inferences are so far from ever being met or approximated that such inferences are never valid. Alternative research methodologies are then briefly discussed.

Keywords: **Comparative Education. Research Methods. Regression Analysis. Causal Inference.**

Introdução

A comparação é uma parte fundamental da ciência e das ciências sociais. O campo da educação comparada foi inicialmente embasado em abordagens culturais e históricas mais qualitativas para fazer comparações (Edwards; Holmes; Van de Graff, 1973; Schriewer; Holmes, 1992). Eu diria que, em geral, isto era verdadeiro em muitas ciências sociais e campos aplicados como a educação até a década de 1950 e 1960, quando os métodos estatísticos começaram a ser largamente aplicados. Uma virada na educação comparada foi a publicação, em 1969, do livro *Toward a Science of Comparative Education*, de Harold Noah e de Max Eckstein. Os autores argumentavam que os métodos culturais e históricos de comparação que tinham sido usados na educação comparada em geral não eram suficientemente científicos ou precisos, sendo necessário aplicar de maneira disseminada os métodos quantitativos que estavam sendo usados em ciências sociais como economia e sociologia. Ainda que atualmente haja um ressurgimento do interesse por métodos qualitativos, mesmo na educação comparada (Bray; Adamson; Mason, 2007), os métodos quantitativos ainda dominam, especialmente na arena política.

O foco deste artigo é a análise de regressão, que constitui o núcleo de uma família de técnicas incluindo análise de trilha, modelagem em equações estruturais, modelo linear hierárquico e outros. A análise de regressão faz todas as comparações de maneira direta ao transformar todas as categorias (países, regiões, raças, gêneros, classes, programas, políticas etc.) em variáveis cujo impacto é medido em resultados de regressão. A análise de regressão talvez seja o método quantitativo mais usado nas ciências sociais, mais especialmente em economia e em sociologia, mas foi utilizada até mesmo em campos como antropologia e história. A análise de regressão na pesquisa (e nos experimentos) em educação ainda é encarada como a abordagem mais objetiva e científica. Constitui a principal base para determinar o impacto da educação e de outras políticas sociais e, como tal, tem enorme influência em quase todas as decisões de políticas públicas.

Este artigo levanta questões fundamentais a respeito da utilidade da análise de regressão para inferências causais. Eu diria que o alcance ou até mesmo a proximidade das condições necessárias para que a análise de regressão produza inferências causais válidas está mais distante do que nunca, de modo que estas inferências nunca são válidas. Esta conclusão desanimadora resulta claramente do exame destas condições no contexto de três exemplos amplamente estudados de análise de regressão aplicada: função de rendimentos, função de produção da educação e função de produção agregadas. Considerando que, na educação comparada, meu campo de especialização é a economia da educação, abordo cada um destes exemplos a partir desta perspectiva. Todavia, penso que minhas conclusões não são particulares ao olhar sobre o impacto da educação ou a estes três exemplos, mas que os problemas subjacentes apresentados são verdadeiros ao fazer inferências causais das análises de regressão sobre outras variáveis e em outros tópicos.

Argumento Geral

Em alguns campos, a análise de regressão é usada como um exercício empírico *ad hoc* para ir além das correlações simples. Muitas vezes os pesquisadores estão interessados no impacto de determinada variável independente sobre determinada variável dependente e usam a análise de regressão como uma maneira de controlar algumas covariáveis. Apesar de comuns, em muitos campos estas sondagens empíricas são reprovadas porque o resultado de interesse particular (o coeficiente sobre a variável independente chave sob exame) dependerá de quais covariáveis são selecionados como *controle*.

Ao invés disso, hoje em dia a maioria dos campos ensina que deve-se ter seriedade em relação à modelagem causal para usar a análise de regressão para inferências causais. Os modelos causais exigem determinadas circunstâncias para assegurar que os coeficientes de regressão sejam estimativas exatas e não-tendenciosas do impacto causal. Embora muitas vezes estas circunstâncias sejam expressas como propriedades de regressão residuais, também podem ser expressas como três condições necessárias para a especificação apropriada de um modelo causal examinando determinada (ou um conjunto de) variável(is) dependente(s):

- Todas as variáveis relevantes estão incluídas no modelo;
- Todas as variáveis são medidas adequadamente; e
- As inter-relações funcionais corretas das variáveis estão especificadas.

Para alcançar uma especificação apropriada, deve-se ter uma teoria muito bem elaborada que permita que estas condições sejam preenchidas². O problema fundamental com análises de regressão é que não temos teorias suficientemente completas em nenhum de nossos campos para especificar adequadamente os modelos causais. Portanto, a literatura sobre aplicação da análise de regressão em geral se transforma em discussões sobre o grau de especificação incorreta e de suas consequências. Infelizmente, a teoria da análise de regressão é muito impiedosa; havendo apenas uma variável omitida, todos os coeficientes de regressão podem ser enviesados em uma medida e em uma direção desconhecidas. Quando os pesquisadores às vezes usam o raciocínio *ad hoc* para inferir a direção do sentido de determinadas variáveis omitidas, o fazem com base em sua correlação potencial com determinada variável independente de interesse incluída. Entretanto, este raciocínio *ad hoc* não é válido. A direção do viés dependerá da intercorrelação da variável omitida com *todas* as variáveis incluídas. O raciocínio *ad hoc* não oferece um indício a respeito de quão enviesados são os coeficientes incluídos.

Mais precisamente, nunca estamos falando sobre o exemplo simples de uma única variável omitida. Defrontamo-nos com falhas múltiplas de todos estes três pressupostos: muitas variáveis são omitidas sempre; sabemos pouco a respeito de como medir melhor as variáveis que conseguimos incluir; e não temos a menor ideia de sua forma fun-

cional. Isto é mais bem ilustrado ao examinar exemplos concretos da literatura referente a análises de regressão, como faço a seguir.

Função de Rendimentos

A função de rendimentos é usada principalmente por economistas e sociólogos para investigar as determinantes de diferenças nos rendimentos. Provavelmente um dos tópicos de estudo que sofrem maior regressão e tem sido especialmente relevante para a economia da educação como fonte de taxa de retorno a estimativas de educação (Blaug, 1976; Psacharopoulos; Patrinos, 2004). Penso que a função de rendimentos é especialmente interessante porque é um dos poucos terrenos onde os cientistas sociais de esquerda e de direita vem competindo, principalmente devido a discussões sobre a segmentação do mercado de trabalho. Em economia, isto era um desafio à ideia neoclássica de que havia apenas um grande mercado de trabalho perfeito no qual o sucesso era determinado pelas características individuais que alguém tivesse como capital humano. Ao contrário, os economistas políticos e outros críticos da história neoclássica enxergavam um mercado de trabalho imperfeito com fraturas (por exemplo, divisões em mercado de trabalho primário e secundário) e estruturas (por exemplo, grandes empresas, sindicatos, sexismo e racismo) que influenciavam imensamente a possibilidade de um indivíduo ser bem-sucedido. Em sociologia, isto era um desafio semelhante à ideia originária da teoria estrutural-funcionalista dominante e de sua teoria do alcance de status derivada que, como a economia, afirmava que o sucesso individual era determinado principalmente por características individuais. Ao invés disto, os sociólogos críticos, muitas vezes compartilhando uma teoria do conflito crítica ao estrutural-funcionalismo, argumentavam, como economistas políticos, que o sucesso no mercado de trabalho era enormemente determinado por fatores estruturais. Cada lado neste debate utilizou a análise de regressão para *provar* seu ponto de vista (Klees; Milton, 1993).

Mais importante ainda é que literalmente há centenas de estudos sobre a função de rendimentos, sendo que cada estudo utiliza desde muito poucas especificações até especificações vastamente diferentes. As três principais condições necessárias para que os coeficientes de regressão de uma função de rendimentos sejam estimadores precisos do verdadeiro impacto causal estão muito longe de serem preenchidas. Primeiramente, todas as variáveis relevantes que podem afetar os rendimentos podem nunca vir a ser incluídas. Nossas teorias literalmente propõem dezenas de variáveis, sendo que decidir quais variáveis serão incluídas em determinado estudo de regressão é novamente idiossincrático. Os exemplos de variáveis que alguns pesquisadores consideram relevantes são: status de saúde, anos de escolaridade, qualidade da escolaridade, tipo de escolaridade, habilidade cognitiva, raça, etnia, religião, status socioeconômico, gênero, status de imigração, status conjugal, viver em união, busca de trabalho, status e diferenciação da ocupação, segmento do mercado de trabalho, características da empre-

sa e da indústria e muito mais. Em segundo lugar, não conhecemos a maneira *correta* de medir a maioria destas variáveis. A medida é *ad hoc* e varia de estudo para estudo. Em terceiro lugar, a inter-relação funcional entre variáveis não é conhecida. Embora seja comum usar o logaritmo natural da renda como a variável dependente, mesmo os economistas neoclássicos admitem que o embasamento para fazê-lo é muito frágil e, na realidade, o que seria necessário é especificar algum conjunto desconhecido de equações simultâneas complexas preenchidas com variáveis sujeitas a interações complexas (Blaug, 1976; Klees; Milton, 1993).

O resultado deste estado as coisas é uma especificação incorreta infinita – por necessidade³. Cada pesquisador possui um rol quase infinito de escolhas na maneira como especifica a função de rendimentos que for estimar. Cada estudo de regressão nunca é uma replicação, mas sim sempre diferente de outros em muitos aspectos. O resultado é que cada estudo de regressão é idiossincrático. Considerando que é relativamente fácil conseguir coeficientes significativos, especialmente com grandes conjuntos de dados, cada um encontra sua variável particular de interesse para ser significativa. Quando existe controvérsia, cada um encontra evidências empíricas que apoiem seu lado no debate. Cada teórico da segmentação encontra segmentos do mercado de trabalho que sejam um fator significativo para determinar os rendimentos e os outros desfechos do mercado de trabalho, embora nenhum economista neoclássico ou sociólogo estrutural funcionalista jamais o faça.

No que concerne à educação, a maioria encontra algum efeito da educação nos rendimentos, relata-o e às vezes o utiliza para estimar uma taxa de retorno. Porém, as especificações alternativas sempre produzem resultados diferentes e, assim, as estimativas são notavelmente instáveis e inconsistentes. Hanushek (1980, p. 240) afirmou que

[...] taxas estimadas de retorno por anos de escolaridade, particularmente em estimativas de regressão (sobre rendimentos), considerando outras diferenças individuais, parecem muito instáveis: mudanças na amostra, mudanças em períodos de tempo, mudanças em especificações precisas de modelo produzem mudanças enormes em taxas estimadas de retorno.

Os impactos estimados da educação sobre os rendimentos e as taxas associadas de retorno são basicamente arbitrários e o resultado de empirismo *ad hoc* corre solto.

Funções de Produção da Educação

Outro uso muito comum da análise de regressão é estimar as chamadas produção da educação ou função insumos-resultados (Levin, 1976; Hanushek, 1986). Em geral, a variável dependente estudada é o escore de um estudante em algum teste de aproveitamento escolar. Novamente, é impossível cumprir as três condições para uma especificação apropriada. Primeiramente, o rol de variáveis independentes em potencial é enorme, incluindo, por exemplo: status socioeconômico, gênero, raça, etnia, idade, dedicação às tarefas de casa, uso de computador em

casa, aprendizagem prévia, habilidade, motivação, aspirações, características dos colegas, nível de graduação do professor, práticas docentes, habilidade do professor, experiência do professor, tamanho da turma, ambiente na escola, características do diretor e políticas curriculares, para mencionar algumas. Em segundo lugar, não existe nenhuma concordância a respeito de como medir a maioria, se não todas, estas variáveis. Em terceiro lugar, mais uma vez as inter-relações funcionais possíveis são inúmeras. Em oposição à formulação linear usualmente executadas, foram propostas formulações de equações recursivas e simultâneas com um rol de termos de interação entre as variáveis independentes, porém foram pouco usadas (Levin, 1976; Hanushek, 1986).

Os economistas da educação, os sociólogos da educação e outros pesquisadores educacionais estimaram centenas destas funções. Mais uma vez, com um rol infinito de escolhas de especificação como este, quase todo estudo é singular e idiossincrático. Hanushek (1979; 1986; 2004) estudou e resumiu, ao longo do tempo, os resultados de estudos como estes. Não surpreende que ele e outros tenham encontrado resultados inconsistentes. Entretanto, ele e a vasta maioria dos pesquisadores quantitativos agarram-se à esperança de que melhorias nos modelos e nos dados podem no fim mostrar alguns resultados claros. Ao contrário, observei a completa indeterminação desta forma de pesquisa integrada às próprias suposições sobre as quais está embasada.

Um uso particularmente destrutivo destas funções é para o assim-chamado pagamento por desempenho para professores. O *valor-adicionado* ao escore do aproveitamento escolar do estudante por professores individuais é determinado por meio da estimativa de uma função de produção educacional, em geral usando apenas algumas variáveis controle, sendo os efeitos do professor determinados por variáveis binárias ou residuais (American Educational Research Association, 2016). O problema, obviamente, é que, com diferentes variáveis controle, professores diferentes são bem ou mal classificados e não existe lógica para escolher determinada especificação e não outra. Mesmo assim, nos EUA os professores estão sendo contratados e dispensados com base nestes resultados completamente espúrios.

Função de Produção Agregada

Embora muitos estudos de economia da educação tenham examinado o impacto dos insumos da educação sobre o aproveitamento escolar e outros tenham focado a conexão entre educação e rendimentos, como um indicador para a produtividade, alguns estudos tentaram examinar mais diretamente a conexão entre educação e produtividade ao examinar o efeito da educação sobre o crescimento econômico, medido pelo PIB. De fato, alguns dos trabalhos mais iniciais sobre o capital humano examinaram a correlação entre níveis de educação ou matrículas escolares em um país e em seu PIB (Bowman, 1966; Blaug, 1970). Entretanto, correlação não é causalidade e estes estudos foram rapidamente descartados porque nem controlavam outras diferenças entre países nem demonstravam qual era a causa e qual era o efeito (Blaug, 1970).

O trabalho inicial mais significativo e ainda extensamente citado que tentou examinar de maneira mais sofisticada a conexão entre educação e PIB foi o de Edward Denison (1961, 1967). Denison enfocou uma determinada forma do que os economistas chamam de “[...] função de produção agregada” (Denison, 1961; 1967). Assim como uma função de rendimentos tenta examinar todas as variáveis que poderiam afetar os rendimentos, as funções de produção examinam mais diretamente todas as variáveis que poderiam afetar os resultados em determinada indústria. Uma função de produção *agregada*, como implica o nome, examina o efeito de insumos sobre o resultado total da produção, isto é, o PIB. Esta abordagem, *em teoria*, poderia contornar a necessidade de pressupor que os rendimentos refletem a produtividade ao examinar diretamente o impacto da educação sobre o resultado. No entanto, o famoso trabalho de Denison *não* fez isto. Ao invés de estimar uma função de produção agregada, *pressupôs* uma função de maneira particular e então usou a associação da educação com rendimentos como a evidência do impacto da educação sobre o PIB, fornecendo com isto nada diferente dos resultados oferecidos pela problemática conexão educação-rendimentos discutida acima. Blaug (1970) ignorou toda esta pesquisa inicial: “Em resumo, aprendemos com as comparações internacionais [de educação e PIB]... que não aprendemos com comparações internacionais” (Blaug, 1970, p. 100).

Tentar conectar diretamente educação com PIB em geral caiu em desuso até o fim das décadas de 1980 e 1990, quando alguns trabalhos na área do que era chamada *nova teoria do crescimento* sinalizaram uma visão mais ampla da contribuição da educação (Romer, 1986; Lucas, 1996; Psacharopoulos; Patrinos, 2004)⁴. Esta visão é teoricamente interessante porque a educação é encarada não apenas como contribuinte para a produtividade do trabalhador, mas como favorecedora do crescimento por meio de uma variedade de mecanismos e exterioridades. Entretanto, empiricamente estas novas direções provaram ser extremamente difíceis de modelar matematicamente. Quase todo pesquisador que tenta estimar estas conexões usa, portanto, um modelo diferente e os resultados são, como seria esperado, tipicamente idiossincráticos, instáveis e inconsistentes (Psacharopoulos; Patrinos, 2004; Stevens; Weale, 2004). Em 1970, Blaug disse que a “Meca da economia da educação se encontra em outro lugar” (Blaug, 1970, p. 100), e penso que isto continua verdadeiro atualmente por razões semelhantes àquelas que discuti para educação e rendimentos e para insumos educacionais sobre resultados.

Conforme eu já disse, os resultados da pesquisa empírica estimando os impactos acima têm sido idiossincráticos, instáveis e inconsistentes. O mesmo é verdadeiro para o impacto da educação sobre o PIB por razões similares. Em primeiro lugar, não existe concordância sobre como medir o estoque ou o fluxo de capital humano em um país. Foram usados vários indicadores, mas, conforme Psacharopoulos e Patrinos (2004) admitem, esta medida pode ser *o ponto o mais fraco* destes estudos: “Dados como estes têm sérios problemas de comparabilidade intertemporal e interpais e existem lacunas nos dados muitas

vezes preenchidas com dados construídos com base em interpolações e extrapolações” (Psacharopoulos; Patrinos, 2004, p. 13-14).

Em segundo lugar, mais de acordo com minhas considerações, conforme Psacharopoulos e Patrinos (2004, p. 15) também admitem: “Os países também diferem em muitos outros aspectos além daqueles medidos pelo estoque físico e de capital humano...” que podem afetar o PIB. As estimativas de funções de produção agregadas literalmente têm usado dezenas de variáveis diferentes como insumos, como clima, latitude, acesso a vias navegáveis, infraestrutura de transporte, desenvolvimento tecnológico, ambiente de investimento, diferenças culturais e políticas, política fiscal e monetária etc. (Stevens; Weale, 2004; Hulten; Issakson, 2007; Hulten, 2009)⁵. Estudos empíricos escolhem de maneira idiossincrática algumas destas variáveis de insumos entre aquelas disponíveis no conjunto de dados sendo usados e sempre omitem muitas outras. Conforme Psacharopoulos e Patrinos (2004, p. 15) admitem outra vez: “Estas variáveis omitidas podem conduzir a margens de erro de centenas por cento em diferenças na trajetória de crescimento econômico entre países”.

Em terceiro lugar, os economistas reconhecem largamente que a forma funcional linear tão comumente usada em estudos de análise de regressão não se aplica a funções de produção agregada. Entretanto, existe um considerável debate acerca de qual forma funcional usar e formas funcionais diferentes produzem estimativas diferentes do impacto da educação (e de todos os outros insumos) sobre o PIB (Stevens; Weale, 2004). Existe até mesmo uma respeitada escola de economia que diz que não há nenhuma base teórica para acreditar que uma função de produção agregada realmente exista. Cada bem e serviço pode ter uma *função de produção*, significando alguma regularidade matemática em como recursos como terra, trabalho, capital e tecnologia se combinam para produzir televisores, iates, políticas de seguridade, hambúrgueres etc. No entanto, como não há nenhum processo físico pelo qual o PIB agregado seja produzido, nem existe, a partir desta perspectiva, alguma maneira de agregar e medir o capital físico, tentar especificar uma função de produção agregada é considerado sem sentido (Cohen; Harcourt, 2003). Guerrien e Gun (2015, p. 100) observam que Paul Samuelson, vencedor do Prêmio Nobel em economia, indicou que as funções de produção agregadas erroneamente oferecem “[...] um teste estatístico de uma identidade contabilística (que, por definição, sempre é verdadeira)”. Defendem a necessidade de “[...] convencer a todos para abandonarem definitivamente as funções de produção agregadas [sic] tanto na teoria como na prática” (Guerrien; Gun, 2015, p. 99) (consulte também Felipe; McCombie, 2013).

Dados estes problemas fundamentais no cumprimento das condições para que a análise de regressão produza estimativas precisas do impacto causal (discutido anteriormente), não surpreende que não sejam encontrados resultados consistentes do impacto da educação sobre o PIB. As revisões desta literatura relatam um confuso rol de escolhas metodológicas idiossincráticas que têm como resultado um confuso rol de resultados diferentes (Stevens; Weale, 2004). Psacharopoulos e Pa-

trinos (2004, p. 15) citam Temple e Voth (1998, p. 1359): “Tentar impor a estrutura de uma função de produção agregada quase certamente é a abordagem equivocada para muitos países em desenvolvimento”. Eu diria que esta abordagem é equivocada para qualquer país⁶.

Deve ser observado que quase todos estes estudos oferecem apenas alguma medida da quantidade de educação, não sua qualidade. Em um estudo recente amplamente citado, Hanushek e Woessmann (2008) tentam solucionar isto acrescentando escores médios de teste PISA do país como um indicador para a qualidade da educação em um país, concluindo que uma diferença de um desvio padrão em escores do teste produz uma taxa de crescimento de 2 pontos percentuais mais elevada de PIB/capita. À luz dos problemas supracitados, acho esta afirmativa completamente irracional e sua recepção acrítica devido à ignorância dos problemas fundamentais com a teoria e o empirismo do capital humano discutidos neste artigo (consulte também Klees, 2016). As medidas de quantidade e qualidade da educação, escolha de outros insumos para controlar e a escolha da forma funcional de Hanushek e Woessmann são todas⁷ idiossincráticas. São apenas uma entre literalmente milhares de especificações alternativas razoáveis de uma função de produção agregada. Especificações diferentes produzirão resultados diferentes⁸.

Discussão

Embora eu tenha abordado os exemplos acima como um economista mais interessado no impacto da educação, os problemas são idênticos ao examinar o impacto de qualquer uma da miríade de variáveis independentes nestas equações. Além disso, no meu entendimento, a impossibilidade de uma especificação apropriada em geral é verdadeira em análises de regressão entre as ciências sociais, seja se estivermos examinando fatores que afetam o status ocupacional, o comportamento eleitoral etc. O problema é que como foi inferido pelas três condições para que as análises da regressão produzam estimativas precisas, não tendenciosas, você precisa investigar um fenômeno que tem regularidades matemáticas subjacentes - e, além disso, você precisa saber quais são. Nenhuma delas parece ser verdadeira. Não tenho motivo para crer que a maneira pela qual múltiplos fatores afetam os rendimentos, o aproveitamento escolar e o PIB tenha qualquer regularidade matemática subjacente entre indivíduos ou países. Mais provavelmente, cada indivíduo ou país têm uma função diferente e que se modifica ao longo do tempo. Mesmo que houvesse alguma constância, os processos são tão complexos que não temos nenhuma ideia de como a função se parece.

Os pesquisadores aceitam que não conhecem a verdadeira função e parecem tratar seus resultados, usualmente de maneira implícita, como uma abordagem suficiente. Mas não há nenhuma base para a crença de que os resultados do que é executado na prática seja qualquer coisa próxima do fenômeno subjacente, ou até mesmo se existe um fenômeno subjacente. Isto parece ser apenas uma ilusão. A maior parte da pesquisa em análise de regressão nem mesmo apoia falsamente regula-

ridades teóricas. Mas você não pode apenas fazer regressão de qualquer coisa que quiser e esperar que os resultados se aproximem da realidade. E mesmo quando os pesquisadores levam muito a sério a necessidade de ter uma estrutura teórica subjacente – como têm, pelo menos em alguma medida, nos exemplos de estudos de rendimentos, de aproveitamento escolar e PIB que usei para ilustrar meu argumento – estão tão distantes das condições necessárias para uma especificação apropriada que não se pode ter nenhuma confiança na validade dos resultados.

Ademais, o que os pesquisadores fazem na prática invalida ainda mais seus resultados. Em teoria, ao usar a análise de regressão, supostamente você começa com uma especificação de modelo completo e depois pega seus dados e faz a estimativa, de uma vez só. Dada a indeterminação da especificação do modelo, ninguém faz isto na prática. Em seu artigo agora clássico, *Let's Take the Con Out of Econometrics*, Leamer (1983, p. 36) descreve a análise de regressão no mundo real e suas consequências:

A arte econométrica da maneira como é praticada no computador... envolve o ajuste de muitos, talvez milhares de modelos estatísticos... Esta busca por um modelo muitas vezes é bem-intencionada, mas não pode haver nenhuma dúvida de que esta busca de especificação invalida as teorias de inferência tradicionais. Os conceitos de não-tendenciosidade, consistência, eficiência, estimativa de probabilidade máxima, de fato, todos os conceitos da teoria tradicional perdem completamente seu sentido pelo tempo que um pesquisador aplicado gasta com o espinheiro do resultado do computador para extrair o único espinho de um modelo do qual mais gostar, aquele que escolher retratar como uma rosa.

Para mim, a pergunta prática transforma-se em se aprendemos algo a partir de todas estas pesquisas? A maioria dos pesquisadores quantitativos diria que aprenderam, mas creio que esta aprendizagem, se for examinada, seria transformada em um subconjunto de estudos feitos a partir de uma perspectiva com a qual o pesquisador concordava. Conforme Leamer (1983, p. 37) escreveu: “Quase ninguém leva a sério as análises de dados. Ou, talvez mais precisamente, quase ninguém leva a sério as análises de dados dos outros” (consulte também Leamer, 2010). Quase ninguém jamais usa a especificação de alguém sem *aperfeiçoá-la*, argumentando explícita ou implicitamente que o estudo anterior estava incorreto.

Estas observações não implicam que, pelo menos dentro de paradigmas, não exista nenhuma aprendizagem cumulativa a partir dos argumentos uns dos outros. Esta aprendizagem de fato ocorre. Entretanto, o argumento aqui sugere que a inferência causal baseada em regressão simplesmente é uma desculpa para teorizar, porém não fornece nenhuma evidência válida para tal. Existe um antigo ditado em economia: *Se você torturar os dados por tempo suficiente, a natureza confessará*. Na realidade, a natureza nunca confessa. Estudos a partir dos três exemplos que escolhi ordenaram a atenção de educadores e le-

gisladores por mais de 50 anos, embora, na realidade, eu acredite que esta abordagem não tem nenhuma validade, não fornecendo nenhuma informação confiável ou mesmo aproximada para ajudar uma alocação sensata de recursos societais.

Os econométristas e outros analistas de regressão reconhecem que há muitas fontes de viés de coeficientes de regressão. Despendem muito tempo com maneiras de corrigir coisas como viés de seleção da amostra e erro de medida - sem muito sucesso, a menos que você esteja disposto a fazer algumas suposições heroicas. Mas estes problemas são menores quando comparados com especificação incorreta desenfreada. Os analistas de regressão tentaram lidar com um problema de especificação incorreta - aquele das variáveis omitidas - por meio do uso de variáveis instrumentais (VIs). Porém, em geral isto requer uma medida precisa de variáveis incluídas e uma especificação correta de forma funcional, nenhuma das quais é verdadeira. As técnicas de variáveis instrumentais fornecem resultados diferentes, dependendo da VI escolhida, além de outros problemas (Heckman; Urzua, 2009; Leamer, 2010). Mais uma vez, estas e outras técnicas (descontinuidade da regressão, diferenças-em-diferenças) requerem pressupostos heroicos para lidar com qualquer aspecto de especificação incorreta (Angrist; Pischke, 2009)⁹.

Creio que, infelizmente, a metodologia da análise de regressão seja um impasse, nada melhor do que alquimia e frenologia, e um dia as pessoas olharão para trás, maravilhadas como pessoas tão inteligentes conseguiam se convencer de outra maneira. Este não é um problema que melhores modelagens, técnicas e dados possam reparar¹⁰.

Alternativas

Não encaro a essência do problema como quantificação. Nem penso que seja fútil tentar procurar causas e consequências de nossas práticas e políticas. Quantificar fenômenos sociais claramente tem seus limites e, no máximo, produz aproximações (Samoff, 1991). Mas tabulações cruzadas e correlações são úteis para sugerir inter-relações. Como é bem sabido, entretanto, qualquer associação encontrada pode ser espúria ou ter uma miríade de explicações alternativas. Por exemplo, *crosstabs* podem revelar que, em média, as mulheres ganham \$0,75 em comparação a \$1,00 ganho por homens. Podemos evidenciar isto um pouco mais ao observar mulheres e homens que trabalham em tempo integral, quando talvez os dados mostrem uma comparação de \$0,80 para \$1,00. Podemos ir mais fundo e examinar a situação de mulheres com curso superior que trabalham em tempo integral comparadas a homens em circunstâncias similares, talvez fornecendo uma comparação de \$0,90 para \$1,00. Os *crosstabs* podem fornecer comparações ainda mais refinadas. Estas comparações, apesar das limitações, fornecem dados reais, descritivos, válidos. Infelizmente, a esperança das ciências sociais de que possamos controlar simultaneamente uma diversidade de fatores como educação, inserção à força de trabalho, discriminação e outros é simplesmente uma ilusão.

O problema é que as relações causais subjacentes a estas associações são tão complexas e tão irregulares que o processo mecânico de análise de regressão não tem nenhuma esperança de evidenciá-las. Uma esperança para os pesquisadores quantitativos que reconhecem os problemas que tenho discutido é o uso da experimentação – sendo a terminologia preferida atualmente ensaios clínicos controlados (ECCs). Supostamente os ECCs contornam os problemas enfrentados pela análise de regressão com o uso de cuidadosos controles físicos e experimentais em vez de estatísticos. A ideia é que, ao fazê-lo, será permitido a alguém observar o efeito de um fator individual, como se um estudante frequentou determinado programa de leitura. Para fazer isto, designa-se aleatoriamente estudantes para um grupo experimental e para um grupo controle, o que, em teoria, permitirá uma firme atribuição de causa e efeito. Feito isto, espera-se que a diferença no aproveitamento entre os grupos seja resultado da frequência ao programa de leitura. Infelizmente, pode ser ou não. Você ainda tem o problema que os processos sociais e pedagógicos são tão complexos, com tantos aspectos a considerar, que, juntamente com algumas dimensões relevantes, o grupo controle e o experimental não serão semelhantes. Isto é, se você olhar de perto todos os fatores potencialmente relevantes, quase sempre os grupos controle são sistematicamente diferentes do grupo experimental e o resultado é que já não temos mais a capacidade de fazer inferências claras. Ao contrário, precisamos usar alguma forma de análise estatística para controlar as diferenças entre os dois grupos. Entretanto, a aplicação de controles estatísticos transforma-se um exercício *ad hoc*, ainda pior do que a abordagem de regressão de modelagem causal. Nesta última, pelo menos existe uma pretensão de desenvolver um modelo completo de variáveis potencialmente intervenientes, enquanto, na primeira, algumas covariáveis são selecionados de maneira muito arbitrária como controle. No fim, não se sabe se as diferenças no aproveitamento escolar são devidas ao programa de leitura ou a outros fatores (Leamer, 2010).

Se estivermos interessados em examinar dados quantitativos, receio que fiquemos presos à discussão de tabulações cruzadas e correlações. Esta é uma perspectiva desanimadora para a maioria dos pesquisadores quantitativos que gastaram anos se tornando virtuosos em análise de dados e encaram as implicações de meu argumento como essencialmente abandonar a empreitada da pesquisa. Felizmente, para muitos de nós, a empreitada da pesquisa está viva e bem, com uma miríade de metodologias alternativas mais qualitativas com as quais investigar nosso mundo educacional e social.

Quando fui para a faculdade, as disciplinas de introdução a métodos de pesquisa muitas vezes se concentravam em análise de regressão ou em exame do delineamento e análise de estudos quantitativos experimentais e quasi-experimentais de Campbell e Stanley (1963). Ainda hoje isto é verdadeiro em determinados campos e departamentos universitários. Entretanto, os últimos 30 anos testemunharam um florescimento de abordagens alternativas aos métodos de pesquisa, especialmente em educação, mas em outros campos também. A

educação tem estado na linha de frente destas mudanças em grande medida, na minha opinião, porque muitas das mudanças foram geradas no campo da avaliação de programas que resultou, em grande parte, de avaliações educacionais que foram encomendadas pelo Congresso dos EUA nas décadas de 1960 e 1970. Muitos daqueles envolvidos no trabalho de campo da avaliação simplesmente descobriram que a abordagem quantitativa à pesquisa e à avaliação não conseguia capturar a experiência dos programas que estavam estudando e valeram-se de outras tradições, como em sociologia ou antropologia, ou inventaram novas abordagens. Em anos subsequentes, estas incursões produziram uma larga lista de métodos alternativos para pesquisa e avaliação (Mertens, 2015).

Durante vários anos, tive a sorte de ensinar em uma disciplina de Introdução a Métodos de Pesquisa em nosso departamento. Embora qualquer agrupamento de métodos seja um tanto arbitrário e sua rotulação sempre seja problemático, a disciplina é dividida em três, respectivamente enfocando métodos quantitativos/positivistas, métodos qualitativos/interpretativos e métodos críticos/transformadores. As comparações são tão essenciais aos últimos dois paradigmas quanto para os métodos quantitativos. Existe uma extensa literatura sobre o debate qualitativo/quantitativo. Alguns afirmam que é demasiada, enquanto outros, com quem eu concordo, afirmam que existem diferenças teóricas fundamentais na perspectiva que precisam ser consideradas (Smith; Hesushius, 1986; Mertens, 2015). De qualquer maneira, fica claro que existem muitas alternativas qualitativas às abordagens quantitativas experimentais e de análise de regressão, incluindo estudo de caso, etnografia, teoria fundamentada nos dados, fenomenologia, narrativa e história oral, para mencionar algumas.

Alternativas metodológicas adicionais são oferecidas por perspectivas críticas/transformadoras que provêm da gama de teorias nas ciências sociais e campos aplicados como a economia política radical, a sociologia crítica, os feminismos, a teoria *queer* e outras enfocando aspectos da marginalização (Klees, 2008). Em geral, estas perspectivas criticam a falta essencial de objetividade da pesquisa positivista/quantitativa e da pesquisa qualitativa/interpretativa, argumentando que não existe pesquisa neutra e que, muitas vezes, estes estudos são realizados em apoio a interesses dominantes. A pesquisa crítica/transformadora assume uma posição explícita de trabalhar no interesse das pessoas marginalizadas. Isto inclui a pesquisa sob a denominação de participativa, ação, feminista, indígena, crítica, etnografia crítica e racial crítica (Denzin; Lincoln; Smith, 2007; Smith, 2012; Mertens, 2015)¹¹.

Os proponentes da pesquisa quantitativa reconhecem que alguns destes métodos alternativos existem, mas em geral, no máximo, os relegam ao reino da geração de ideias, não ao processo científico de construção do conhecimento do mundo social. Ao contrário, muitos proponentes de métodos alternativos defendem que são tanto ou mais válidos, confiáveis e generalizáveis do que os quantitativos¹². Por exemplo, Miles e Huberman (1994, p. 434) chegam a afirmar:

Os estudos qualitativos... são especialmente bem adequados para descobrir relações causais; examinam direta e longitudinalmente os processos locais subjacentes a uma série temporal de eventos e estados, mostrando como levaram a desfechos específicos, e descartam hipóteses rivais. De fato, entramos em uma caixa-preta; conseguimos compreender não apenas determinada coisa aconteceu, mas como e por quê aconteceu.

Similarmente, fortes argumentos são feitos em favor da transferibilidade e a generalizabilidade da pesquisa qualitativa e crítica (Donmoyer, 1990; Mertens, 2015).

Conclusões

Durante muitos anos, um colega sociólogo da educação e eu oferecemos uma disciplina de laboratório de regressão. Utilizamos um bom conjunto de dados nacionais e fizemos com que os estudantes passassem o semestre executando especificações alternativas de funções de produção da educação. A cada aula, os grupos vinham e explicavam suas especificações e seus resultados. Conforme era esperado, diferentes especificações de variáveis incluídas, decisões sobre como medir variáveis e formas funcionais produziram resultados substancialmente diferentes. Foi solicitado a cada grupo que explicasse seus resultados como se tivessem sido redigidos para um artigo de periódico. Meu colega sempre costumava comentar sobre a explicação de um grupo referente aos seus resultados da seguinte maneira: *Faz sentido*. E sempre fazia. Da mesma maneira que os artigos na literatura que revisei acima. Sempre podemos identificar o sentido de nossos resultados. Quando executamos regressões, paramos de fazer os muitos ajustes às nossas regressões - que sempre devem ser feitas - quando obtemos resultados que façam sentido para nós. Contudo, pensando nesta literatura como um todo, simplesmente resultam em achados divergentes, todos embasados em especificações alternativas razoáveis de seus modelos de equação de regressão - pelo menos para alguns.

Em conclusão, gostaria de dizer que estou equivocado em meu argumento neste artigo. Seria útil se a roupa não tão nova do imperador fosse mais do que a nudez que os pesquisadores parecem evitar olhar muito de perto. Infelizmente, a teoria e a prática parecem indicar fortemente de outra maneira. As condições teóricas para que a análise da regressão *funcione* nunca estão perto de serem atendidas. E, na prática, as aplicações da análise de regressão parecem resultar em debates intermináveis porque as especificações são tão vagas que os pesquisadores parecem conseguir usar esta família de técnicas para *comprovar* quase qualquer coisa que quiserem. No entanto, quando não conseguimos encontrar as simples regularidades causa-efeito que os analistas de regressão gostariam de descobrir, no mínimo ainda existem muitos métodos alternativos para investigar e fazer comparações em nosso mundo educacional e social.

Tradução do original inglês de Ananyr Porto Fajardo

Notas

- 1 Este artigo é derivado de outro publicado em *Real World Economics Review*. Eu gostaria de agradecer a Jim Cobbe pelos comentários a uma das versões. Quero agradecer especialmente a Sande Milton por seus *insights* e longa colaboração comigo em relação a este tópico.
- 2 As técnicas estatísticas para contornar estas técnicas são discutidas brevemente na conclusão. Resumidamente, o problema é que não conseguem lidar com especificação incorreta desenfreada.
- 3 Imagine que você coletou dados sobre os rendimentos de 100 pessoas aleatoriamente selecionadas. Imagine delinear as literalmente dezenas de fatores que explicariam por que tinham rendimentos diferentes, sendo suas diferenças educacionais apenas um dos fatores. Apesar das técnicas estatísticas sofisticadas, é simplesmente impossível isolar com precisão o impacto da educação ou outra variável de todos os outros fatores.
- 4 Os sociólogos estudam periodicamente os determinantes do PIB, mas em geral sem nenhuma pretensão de justificativa teórica para uma função de produção agregada (por exemplo, Kentor, 1998).
- 5 Tan (2014, p. 426) afirma que “[...] fatores sociais, políticos, institucionais e culturais precisam ser levados em conta quando o impacto da educação” sobre o crescimento econômico for estimado.
- 6 Levine e Renelt (1992, p. 942) discutem a “[...] variedade de políticas econômicas e fatores políticos e institucionais” que precisam ser levados em conta em regressões do PIB: “Como foi descoberto que mais de 50 variáveis estão significativamente correlacionadas com o crescimento em pelo menos uma regressão, os leitores pode ficar inseguros quanto à confiança que devem ter nos achados de qualquer estudo... Descobrimos que apenas poucos achados conseguem suportar alterações leves na lista de variáveis explicativas”. Eu acrescentaria que as alterações do mundo real estão longe de serem leves. Consulte também o artigo de Sala-i-Martin (1997) sobre o assunto, intitulado *I Just Ran Two Million Regressions*.
- 7 Hanushek e Woessmann (2008) controlam apenas duas das literalmente dezenas de variáveis que poderiam ter sido incluídas, sendo que estas duas variáveis refletiam a ideologia neoliberal de desenvolvimento (abertura ao comércio e proteção aos direitos de propriedade). Controles diferentes produziriam, é claro, estimativas muito diferentes do impacto da *qualidade* da educação.
- 8 Hanushek e Woessmann (2015) então, a partir destas estimativas inválidas, fazem projeções país por país para décadas no futuro do que o PIB seria se os resultados do PISA melhorassem. Seus resultados não têm nenhuma validade, dependendo de literalmente centenas de suposições e vínculos causais completamente frágeis.
- 9 Quando muito, os desvios da especificação apropriada devem ser muito bem-comportados para que as ações corretivas funcionem – e nunca o são. Outra abordagem para lidar com especificação incorreta é a análise de sensibilidade, porém é improvável que mesmo sua forma extrema, defendida por Leamer (1983; 2010), de que seria necessário levar em conta amplas variações em especificação, revele quaisquer regularidades robustas, como fica evidente a partir da maior

parte da literatura empírica (consulte também McAleer; Pagan; Volker, 1985 e Saltelli, 2008).

10 Chego à conclusão oposta ao artigo de Hendry (1980), *Econometrics – Alchemy or Science?* É interessante observar que a litania de Hendry a respeito da crítica de Keynes sobre métodos estatísticos é análoga à minha. De acordo com Hendry: “Keynes chegou perto de afirmar que nenhuma teoria econômica é jamais testável” (Hendry, 1980, p. 396) (consulte também Pratten, 2005).

11 Estas classificações não são nítidas. Por exemplo, você pode fazer teoria fundamentada nos dados ou pesquisa feminista a partir de uma perspectiva positivista, interpretativa ou crítica. Para confundir ainda mais, pesquisadores de paradigma crítico podem se valer de qualquer método, inclusive quantitativos.

12 Devido aos problemas com o paradigma positivista, existe considerável literatura examinando critérios alternativos para uma boa pesquisa de acordo com os paradigmas interpretativo e crítico (Mertens, 2015).

Referências

AMERICAN Educational Research Association (AERA). AERA Statement on the Use of Value-Added Models (VAM) for the Evaluation of Educators and Educator Preparation Programs. *Educational Researcher*, California, v. 44, n. 8, p. 448-452, jun. 2016. [2015].

ANGRIST, Joshua David; PISCHKE, Jorn-Steffen. **Mostly Harmless Econometrics: an empiricist’s guide**. Princeton: Princeton University Press, 2009.

BLAUG, Mark. **An Introduction to the Economics of Education**. London: Penguin, 1970.

BLAUG, Mark. The Empirical Status of Human Capital Theory: a slightly jaundiced survey. *Journal of Economic Literature*, Pittsburgh, v. 14, n. 3, p. 827-855, Sep. 1976.

BOWMAN, Mary Jean. The Human Investment Revolution in Economic Thought. *Sociology of Education*, California, v. 39, n. 2, p. 111-137, 1966.

BRAY, Mark; ADAMSON, Bob; MASON, Mark. **Comparative Education Research: approaches and methods**. 2. ed. New York: Springer, 2007.

COHEN, Avi; HARCOURT, Geoffrey Colin. Retrospectives: whatever happened to the Cambridge capital theory controversies? *Journal of Economic Perspectives*, Pittsburgh, v. 17, n. 1, p. 199–214, 2003.

DENISON, Edward Fulton. **The Sources of Economic Growth in the United States and the Alternatives Before Us**. New York: Committee for Economic Development, 1961.

DENISON, Edward Fulton. **Why Growth Rates Differ?** Washington: Brookings, 1967.

DENZIN, Norman; LINCOLN, Yvonna; SMITH, Linda. **Handbook of Critical and Indigenous Methodologies**. Thousand Oaks, California: Sage, 2007.

DONMOYER, Robert. Generalizability and the Single-Case Study. In: EISNER, Elliot; PESHKIN, Alan. **Qualitative Inquiry in Education**. New York: Teachers College Press, 1990. P. 175-200.

EDWARDS, Reginald; HOLMES, Brian; VAN DE GRAFF, John. **Relevant Methods in Comparative Education**. Hamburg: UNESCO Institute of Education, 1973. Available at: <<http://unesdoc.unesco.org/images/0000/000088/008805eo.pdf>>. Accessed on: 01 Feb. 2014.

- FELIPE, Jesus; MCCOMBIE, John. **The Aggregate Production Function and the Measurement of Technical Change**: not even wrong. Cheltenham: Edward Elgar, 2013.
- GUERRIEN, Bernard; GUN, Ozgur. Putting an End to the Aggregate Function of Production. **Real-World Economics Review**, online, v. 73, p. 99-109, 2015.
- HANUSHEK, Eric. Conceptual and Empirical Issues in the Estimation of Educational Production Functions. **Journal of Human Resources**, Madison, v. 14, n. 3, p. 351-388, 1979.
- HANUSHEK, Eric. Alternative Models of Earnings determination and Labor Market Structures. **Journal of Human Resources**, Madison, v. 16, p. 238-259, 1980.
- HANUSHEK, Eric. The Economics of Schooling: production and efficiency in public schools. **Journal of Economic Literature**, Madison, v. 24, n. 3, p. 1141-1177, 1986.
- HANUSHEK, Eric. What If There Are No 'Best Practices'? **Scottish Journal of Political Economy**, Aberdeen, v. 51, n. 2, p. 156-72, 2004.
- HANUSHEK, Eric; WOESSMANN, Ludger. The Role of Cognitive Skills in Economic Development. **Journal of Economic Literature**, Madison, v. 46, n. 3, p. 607-668, 2008.
- HANUSHEK, Eric; WOESSMANN, Ludger. **Universal Basic Skills**: what countries stand to gain. Paris: OECD, 2015.
- HECKMAN, James; URZUA, Sergio. Comparing IV with Structural Models: what simple IV can and cannot identify. **Journal of Econometrics**, Amsterdam, v. 156, n. 1, p. 27-37, 2009.
- HENDRY, David. Econometrics: Aachemy or science? **Economica**, London, v. 47, n. 188, p. 387-406, 1980.
- HULTEN, Charles. Growth Accounting. **National Bureau of Economic Research Working Paper**, Washington, v. 15341, Sep. 2009.
- HULTEN, Charles; ISAKSSON, Anders. Why Development Levels Differ: the sources of differential economic growth in a panel of high and low income countries. **National Bureau of Economic Research Working Paper**, Washington, v. 13469, Oct. 2007.
- KENTOR, Jeffrey. The Long-Term Effects of Foreign Investment Dependence on Economic Growth, 1940-1990. **American Journal of Sociology**, Chicago, v. 103, n. 4, p. 1024-1046, 1998.
- KLEES, Steven. Reflections on Theory, Method, and Practice in Comparative and International Education. **Comparative Education Review**, Chicago, v. 52, n. 3, p. 301-328, aug. 2008.
- KLEES, Steven. Human Capital Theory and Rates of Return: brilliant ideas or ideological dead ends? **Comparative Education Review**, Chicago, v. 60, n. 4, p. 644-672, Nov. 2016.
- KLEES, Steven; MILTON, Sande. Inferences From Regression Analysis: the case of earnings functions. In: COMPARATIVE AND INTERNATIONAL EDUCATION SOCIETY ANNUAL MEETING, 1993, Kingston (Jamaica). **Paper Presented...** Kingston: 1993.
- LEAMER, Edward. Let's Take the Con Out of Econometrics. **American Economic Review**, Pittsburgh, v. 73, p. 31-43, mar. 1983.
- LEAMER, Edward. Tantalus on the Road to Asymptopia. **Journal of Economic Perspectives**, Pittsburgh, v. 24, n. 2, p. 31-46, 2010.
- LEVIN, Henry. Concepts of Economic Efficiency and Educational Production. In: FROMKIN, Joseph. **Education as an Industry**, Cambridge: Ballinger, 1976.

- LEVINE, Ross; RENELT, David. A Sensitivity Analysis of Cross-Country Growth Regressions. **American Economic Review**, Nashville, v. 82, n. 4, p. 942-963, 1992.
- LUCAS, Robert. On the Mechanics of Economic Development. In: GROSSMAN, Gene. **Economic Growth: theory and evidence**, v. 1. P. 284-323. Cheltenham: Elgar, 1996.
- MCALCER, Michael; PAGAN, Adrain; VOLKER, Paul. What Will Take the Con Out of Econometrics. **American Economic Review**, Pittsburgh, v. 75, n. 3, p. 293-307, 1985.
- MERTENS, Donna. **Research Methods in Education and Psychology: integrating diversity with quantitative and qualitative approaches**. 4. ed. Thousand Oaks: Sage, 2015.
- MILES, Matthew; HUBERMAN, Michael. **Qualitative Data Analysis**. 2. ed. Newbury Park: Sage, 1994.
- NOAH, Harold; ECKSTEIN, Max. **Toward a Science of Comparative Education**. New York: Macmillan, 1969.
- PRATTEN, Stephen. Economics as Progress: the LSE approach to econometric modelling and critical realism as programmes for research. **Cambridge Journal of Economics**, Cambridge, v. 29, n. 2, p. 179-205, 2005.
- PSACHAROPOULOS, George; PATRINOS, Harry. Human Capital and Rates of Return. In: JOHNES, Geraint; JOHNES, Jill. **International Handbook on the Economics of Education**. Northampton: Edward Elgar, 2004. P. 1-57.
- ROMER, Paul. Increasing Returns and Long-Run Growth. **Journal of Political Economy**, Chicago, v. 94, n. 5, p. 1002-1037, 1986.
- SALA-I-MARTIN, Xavier. I Just Ran Two Million Regressions. **American Economic Review**, Pittsburgh, v. 87, n. 2, p. 178-183, 1997.
- SALTELLI, Andrea et al. **Global Sensitivity Analysis: the primer**. Chichester: Wiley, 2008.
- SAMOFF, Joel. The Façade of Precision in Education Data and Statistics: a troubling example from Tanzania. **Journal of Modern African Studies**, Cambridge, v. 29, n. 4, p. 669-689, 1991.
- SCHRIEWER, Jgen; HOLMES, Brian. **Theories and Methods in Comparative Education**. 3. ed. New York: Peter Lang, 1992.
- SMITH, Linda. **Decolonizing Methodologies: research and indigenous peoples**. New York: Zed, 2012.
- SMITH, John; HESHUSIUS, Lous. Closing Down the Conversation: the end of the quantitative-qualitative debate among educational inquirers. **Educational Researcher**, California, v. 15, n. 1, p. 4-13, 1986.
- STEVENS, Philip; WEALE, Martin. Education and Economic Growth. In: JOHNES, Geraint; JOHNES, Jill. **International Handbook on the Economics of Education**. Northampton: Edward Elgar, 2004. P. 164-188.
- TAN, Emrullah. Human Capital Theory: a holistic criticism. **Review of Educational Research**, California, v. 84, n. 3, p. 411-445, Sep. 2014.

Steven J. Klees é Professor de Educação Internacional e Comparada na University of Maryland. cursou seu Doutorado na Stanford University e lecionou na Cornell University, Stanford University, Florida State University e Universidade Federal do Rio Grande do Norte, Brasil. Foi Pesquisador Fullbright por duas vezes na Universidade Federal da Bahia, Brasil. O trabalho do professor Klees examina a economia política da educação e do desenvolvimento, com interesses específicos de pesquisa em globalização, neoliberalismo e educação. E-mail: sklees@umd.edu