

UNIVERSIDADE ESTADUAL DE MARINGÁ

FLORIAN IMMANUEL SCHUMACHER

## **A função de capital humano**

*Externalidades setoriais*

Maringá - PR

2010

FLORIAN IMMANUEL SCHUMACHER

## **A função de capital humano**

*Externalidades setoriais*

Dissertação submetida ao Programa de Pós-Graduação em Ciências Econômicas da UEM, como parte dos requisitos necessários para a obtenção do título de Mestre em Economia.

Orientador: Prof. Dr. Joilson Dias

Maringá - PR

2010

Dados Internacionais de Catalogação-na-Publicação (CIP)  
(Biblioteca Central - UEM, Maringá – PR., Brasil)

S392f Schumacher, Florian Immanuel  
A função de capital humano : Externalidades  
setoriais / Florian Immanuel Schumacher. -- Maringá,  
2010.

62 f.

Orientador : Prof. Dr. Joilson Dias.  
Dissertação (mestrado) - Universidade Estadual  
de Maringá, Programa de Pós-Graduação em Ciências  
Econômicas, 2010.

1. Economia - Capital humano. 2. Economia -  
Externalidade de conhecimento. 3. Economia - Retorno  
de escala. 4. Economia - Crescimento econômico -  
Retorno da escolaridade. I. Dias, Joilson, orient.  
II. Universidade Estadual de Maringá. Programa de  
Pós-Graduação em Ciências Econômicas. III. Título.

CDD 21.ed. 338.900981

FLORIAN IMMANUEL SCHUMACHER

## A função de capital humano

*Externalidades setoriais*

Dissertação apresentada como requisito parcial para a obtenção do título de Mestre em Economia, do Programa de Pós-Graduação em Ciências Econômicas da UEM, sob apreciação da seguinte banca examinadora:

Aprovado em \_\_\_\_ de \_\_\_\_\_ de 2010

---

Joilson Dias

Doutor em Economia - UEM

---

Maurício Vaz Lobo Bittencourt

Doutor em Economia - UFPR

---

Maria Helena Ambrósio Dias

Doutora em Economia - UEM

# Agradecimentos

Aos meus pais, pela vida e pelo apoio incondicional.

Aos professores do departamento de Economia, pelos seus ensinamentos. Em especial, aos professores Joilson Dias e Maria Helena Ambrósio Dias, que também me orientaram, e aos professores Edinaldo Tebaldi Maurício Vaz Lobo Bittencourt, pelas contribuições e críticas ao presente trabalho.

À minha família de afinidade, por uma infinidade de coisas.

Aos meus amigos, pelos momentos compartilhados.

À companhia especial da qual pude desfrutar durante o último ano do mestrado, e a qual tornou o percurso muito mais alegre e prazeroso.

Aos colegas do Mestrado, pelas discussões e pela companhia.

À secretária do Programa de Pós-Graduação em Economia, sempre muito prestativa.

A todos que de alguma forma contribuíram para a confecção deste trabalho.

## Resumo

O objetivo desta pesquisa é avaliar se existem efeitos de espraiamento do conhecimento na economia brasileira, conforme proposto por Lucas (1988), Romer (1990) e Acemoglu (1996). O efeito espraiamento é verificado no setor de serviços de apoio às empresas, constituído pelos seguintes subsetores: telecomunicações, serviços de consultoria e pesquisa. A hipótese consiste em verificar se estes setores, ao empregar pessoas com escolaridade em nível de mestrado e superior a este, causam um aumento de produtividade nos demais setores. Estimativas dos retornos à escolarização por grandes setores econômicos são realizadas usando dados da PNAD (2008). O resultado da conjectura proposta mostrou-se positivo e significativo, com um impacto de 3% a 4% sobre o salário real. Em adição, obteve-se evidências sobre a existência de retornos crescentes à educação no Brasil provavelmente devido ao efeito espraiamento, uma vez que estes retornos estão presentes em maior ou menor grau em todos os grandes setores da economia brasileira.

**Palavras-chaves:** Capital humano; externalidades do conhecimento; retornos de escala; desenvolvimento econômico.

## Abstract

The aim of this research is to evaluate whether there are knowledge spillover effects in the Brazilian economy, as proposed by Lucas (1988), Romer (1990) and Acemoglu (1996). The spillover effect is verified in the business services sector, which consists of the following sub-sectors: telecommunications, consulting services and research. The hypothesis is to verify if the employment by these industries of people educated at the masters level and above causes a productivity increase in other sectors. We estimate the returns to education in major market sectors using data from the Brazilian household survey PNAD (2008). The result of the proposed conjecture was positive and significant, with an impact of 3% to 4% over the real wage. In addition, we obtain evidence for the existence of increasing returns to education in Brazil, probably due to the spillover effect, since these returns are present in greater or lesser degree in all major sectors of the Brazilian economy.

**Keywords:** Human capital; knowledge externalities; returns to scale; economic development.

## Lista de Tabelas

5.1	Estatísticas descritivas . . . . .	32
5.2	Definição das variáveis de setor . . . . .	34
6.1	Regressão do logaritmo do salário . . . . .	36
6.2	Retornos de escala, % . . . . .	37
6.3	Retornos setoriais à escolarização, % . . . . .	38

# Sumário

Lista de Tabelas	4
<b>1 Introdução</b>	<b>6</b>
<b>2 Modelos de crescimento</b>	<b>10</b>
2.1 O modelo neoclássico . . . . .	10
2.2 Modelos de crescimento endógeno . . . . .	14
2.3 Externalidades nos modelos de crescimento . . . . .	17
<b>3 Evidências empíricas</b>	<b>22</b>
3.1 Externalidades do capital humano . . . . .	22
3.2 Retornos da escolarização . . . . .	24
<b>4 Mensuração do capital humano</b>	<b>26</b>
<b>5 Procedimentos metodológicos</b>	<b>29</b>
5.1 Implementação empírica . . . . .	29
5.2 Dados . . . . .	31
<b>6 Resultados</b>	<b>35</b>
<b>7 Considerações finais</b>	<b>39</b>
<b>Referências</b>	<b>41</b>
<b>A Modelo de crescimento de Solow (1956)</b>	<b>47</b>
<b>B Modelo de crescimento de Romer (1990)</b>	<b>51</b>
<b>C Rotina do Stata</b>	<b>53</b>

# 1 Introdução

Um marco na evolução da teoria do crescimento econômico foi o surgimento dos modelos de crescimento endógeno, se destacando, em particular, os trabalhos de Romer (1986, 1990) e Lucas (1988). Intrigados com a não-convergência de renda entre países, os proponentes das teorias de crescimento endógeno enfatizam a possibilidade de influenciar a taxa de progresso tecnológico por meio de políticas. Admite-se a existência de externalidades da acumulação de capital (físico ou humano). As externalidades positivas da acumulação de capital podem ser decorrentes, entre outros, do estoque médio de capital humano (LUCAS, 1988) ou da geração de conhecimento (ROMER, 1990). Em ambos os casos, as externalidades são *modeladas* de modo a gerar retornos não-decrescentes para algum fator e assim permitir a não-convergência de renda entre países no longo prazo. Em Lucas, o estoque médio de capital humano exerce um efeito positivo sobre a produtividade individual. Em Romer, o conhecimento, uma vez gerado no setor de pesquisa, se difunde pela economia pelo fato de ser um bem não-rival e apenas parcialmente excludível, resultando em ganhos de produtividade em todos os setores.

Apesar da atratividade teórica das externalidades nas teorias de crescimento, o mecanismo pelo qual estas externalidades se originam permanece em grande parte uma incógnita. Uma das primeiras tentativas de fornecer um microfundamento para o surgimento das externalidades é o artigo de Acemoglu (1996). De acordo com Acemoglu, externalidades que levem a retornos não-decrescentes só são possíveis com um mercado de trabalho sujeito a um equilíbrio não-walrasiano. No modelo de Acemoglu, o salário não resulta da maximização do lucro, mas é uma proporção do produto, e o encontro entre firmas e trabalhadores ocorre por um processo de pareamento (*matching*).

Mesmo que exista alguma evidência empírica sugerindo que os efeitos externos do capital humano sejam pequenos (ACEMOGLU; ANGRIST, 2001; DUFLO, 2004; CICCONE; PERI, 2006), os estudos realizados até agora têm abordado principalmente os efeitos sobre mercados de trabalho locais, tais como cidades ou, no caso de Acemoglu e Angrist (2001), estados. Poucos pesquisadores estimaram esses efeitos para níveis mais agregados, tais como setores. Os poucos que o fizeram encontraram (i) que graus mais

altos de especialização e concentração setorial estão associados a ganhos salariais (WHEATON; LEWIS, 2002), (ii) que economias de localização e externalidades de capital humano parecem ser fenômenos bastante distintos (WHEELER, 2007), (iii) que o efeito geográfico de espraiamento do conhecimento é limitado (JAFFE; TRAJTENBERG; HENDERSON, 1993; ROSENTHAL; STRANGE, 2008; FU, 2007), (iv) que os retornos à acumulação de capital humano são mais elevados na indústria do que nos serviços e que trabalhadores altamente qualificados tendem a se beneficiar dos efeitos intrasetoriais de espraiamento do capital humano, quanto trabalhadores não altamente qualificados tendem a se beneficiar sobretudo das externalidades *pecuniárias*<sup>1</sup> que surgem entre setores (HEUERMANN, 2009). Em suma, a evidência existente sugere que as externalidades do capital humano são mais relevantes no âmbito setorial do que geográfico.

O objetivo desta pesquisa é avaliar se existem efeitos de espraiamento do conhecimento na economia brasileira, conforme proposto por Lucas (1988), Romer (1990) e Acemoglu (1996). São verificados dois tipos de efeito de espraiamento, utilizando dados da PNAD (2008). O primeiro é o ocasionado pelas firmas que possuem como principal produto de venda a informação e o conhecimento, e, assim, obtêm lucro difundindo conhecimento. As firmas com essa característica podem ser agrupadas em um único setor observável: o setor de serviços de apoio às empresas. Para testar o efeito desse setor sobre a produtividade dos demais setores, incluímos numa equação de salário minceriana a fração estadual de pessoas empregadas nos seguintes setores definidos pela CNAE Domiciliar: telecomunicações; pesquisa e desenvolvimento; atividades jurídicas, de contabilidade, e de pesquisas de mercado e opinião pública; seleção, agenciamento e locação de mão-de-obra. O setor de telecomunicações melhora a produtividade das firmas pela redução de custos, pela maior interação entre as pessoas e a criação de sinergias, e pela agilidade do fluxo de informações. O setor de pesquisa e desenvolvimento está ligado diretamente à geração de novos produtos e serviços, cuja mercantilização permite obter ganhos de monopólio. As atividades jurídicas e de contabilidade reduzem os custos de transação impostos pela legislação e pelas relações existentes entre os agentes. As atividades de pesquisa de mercado e opinião pública permitem a tomada de decisões mais acertadas e a redução de recursos desperdiçados, estreitando o foco e aumentando a eficiência das ações das firmas. A seleção, o agenciamento e a locação de mão-de-obra, por sua vez, impactam diretamente

---

<sup>1</sup>As externalidades pecuniárias resultam de alterações nos preços relativos, e são diferenciadas das externalidades tecnológicas, as quais resultam de alterações nas dotações de recursos.

sobre o processo de pareamento entre trabalhadores e firmas, elevando a produtividade de ambas as partes bem como gerando a externalidade pecuniária de Acemoglu (1996). Como a amostra utilizada é transversal, foi verificado o efeito por unidade federativa.

O segundo efeito de espraiamento é o de Romer (1990). Em seu modelo, as externalidades do conhecimento necessariamente surgem no setor de pesquisa, porque é apenas nele que ocorre o processo de invenção e aprendizagem. Assumindo que na realidade esse tipo de externalidade surja onde quer que ocorra a acumulação de conhecimento, poderíamos verificar efeitos de espraiamento do conhecimento em universidades, institutos de pesquisa, departamentos de pesquisa e desenvolvimento de firmas e, adicionalmente, em todas as atividades econômicas intensivas em capital humano. Se parte dos ganhos monopolistas das firmas<sup>2</sup> forem transferidos aos empregados, esses últimos deveriam ser recompensados com salários mais elevados em setores com maior concentração de pessoas altamente qualificadas.<sup>3</sup> Isto nos permite aproximar a externalidade de Romer pela proporção setorial de pessoas com título de mestre ou superior a este.

Devido à natureza não-rival do conhecimento, os efeitos de espraiamento do conhecimento estão intrinsecamente relacionados a não-convexidade da função de produção agregada, isto é, retornos crescentes de escala. Estes últimos, porém, não são facilmente verificados na prática. Na maioria dos países, observam-se retornos constantes ou decrescentes à acumulação de capital (seja este físico ou humano). Mas, como as estimativas indicam que os retornos à escolarização no Brasil sejam elevados e apresentem retornos crescentes de escala (LAM; SCHOENI, 1993; RAMOS; VIEIRA, 1996; HOFFMANN; SIMÃO, 2005; SALVATO; SILVA, 2008; MONTEIRO; DIAS; DIAS, 2009), optou-se por realizar estimativas dos retornos de escala da escolaridade das pessoas empregadas nos grandes setores econômicos no Brasil.

Além desta introdução, o trabalho consiste em 6 seções. Após revisar brevemente os modelos de crescimento econômico mais relevantes para o objetivo deste trabalho, são apresentados em seguida algumas evidências empíricas sobre a existência de externalidades da acumulação de capital humano e conhecimento, bem como a respeito dos retornos de escolarização no Brasil. A seção quatro trata da mensuração do capital humano e a

---

<sup>2</sup>Apesar de todas essas externalidades serem *a priori* igualmente importantes e interessantes, Romer mostra que apenas as firmas são capazes de internalizar uma parcela substancial delas, quando atuam sob um ambiente de competição monopolista.

<sup>3</sup>Acemoglu (1996) mostra que em um mercado de trabalho em equilíbrio não-walrasiano o salário é uma proporção do produto.

---

seção cinco apresenta os procedimentos metodológicos. As duas últimas seções contêm, respectivamente, os resultados e as considerações finais.

## 2 Modelos de crescimento

A teorização neoclássica a respeito do efeito do progresso tecnológico sobre o crescimento econômico passa por dois momentos principais. Primeiro, o nível tecnológico é modelado como a eficiência de transformação dos insumos capital e trabalho em produto. Neste sentido, inclui não apenas aspectos técnicos de transformação, mas também condições mais gerais que determinam a produção, tal como atuações de política econômica, eventos climáticos, determinantes sociais das relações de trabalho etc. O segundo momento ocorre quando se busca endogeneizar a mudança na eficiência de produção, reconhecendo a possibilidade de influenciar deliberadamente a taxa de progresso técnico. Este reconhecimento está ligado à percepção de que determinadas atividades podem ter efeitos distintos sobre o progresso tecnológico. Passa-se da esfera agregada ou nacional para a esfera setorial, enfatizando as potenciais contribuições de determinados setores para o crescimento econômico de longo prazo. Em particular, o setor educacional e o setor de pesquisa e desenvolvimento ganham destaque, por seu potencial de provocar mudanças na taxa de crescimento do progresso técnico.

A seguir apresentamos o modelo básico correspondente ao primeiro momento da teoria, enquanto na seção subsequente será detalhado um dos diversos modelos que representam o segundo momento teórico.

### 2.1 O modelo neoclássico

A *moderna* teoria de crescimento econômico (isto é, a neoclássica) tem a sua origem mais direta com Solow (1956) e Swan (1956), mas assumiu a sua forma atual com o trabalho complementar de Cass (1965) e Koopmans (1965) que adicionaram ao modelo Solow-Swan um tratamento adequado da otimização intertemporal das famílias (BARRO; SALA-I-MARTIN, 2004). Mostraremos aqui somente o modelo de Solow mais simples, sem a otimização de famílias, porque isto não afeta os resultados principais que queremos sublinhar.

A característica principal do modelo Solow-Swan é o formato da função de produção,

a qual assume retornos constantes de escala, retornos decrescentes para cada insumo, e uma elasticidade de substituição positiva e suave entre os insumos. Este tipo de função de produção, combinada com taxas de poupança e de fertilidade constantes, dá origem a um modelo de equilíbrio geral que prediz convergência condicional de renda per capita entre países (BARRO; SALA-I-MARTIN, 2004). A convergência condicional é o processo de convergência de renda per capita entre países com condições iniciais similares, isto é, instituições, dotações de fatores e preferências intertemporais similares.

A convergência condicional de renda per capita é um resultado direto da hipótese de retornos decrescentes do capital. Outra implicação é que sem que haja um contínuo progresso tecnológico – o qual é exógeno no modelo – o crescimento da renda per capita deve cessar em algum momento. E este é o ponto mais fraco da teoria porque as taxas de crescimento seculares da renda per capita de uma variedade de países não têm mostrado uma tendência a desacelerar. De acordo com Barro e Sala-I-Martin (2004), além deste problema também têm surgido evidência contra a hipótese de taxas de poupança e de natalidade constantes. Mas, a implicação do modelo que encontra evidência empírica favorável é a convergência condicional de renda per capita, isto é, entre velhos membros da OCDE, por exemplo.

A estrutura básica da economia modelada por Solow-Swan é elucidada por Barro e Sala-I-Martin (2004, p. 23):

First, households (or families) own the inputs and assets of the economy, including ownership rights in firms, and choose the fractions of their income to consume and save. Each household determines how many children to have, whether to join the labor force, and how much to work. Second, firms hire inputs, such as capital and labor, and use them to produce goods that they sell to households or other firms. Firms have access to a technology that allows them to transform inputs into output. Third, markets exist on which firms sell goods to households or other firms and on which households sell the inputs to firms. The quantities demanded and supplied determine the relative prices of the inputs and the produced goods.

Uma boa apresentação formal do modelo Solow-Swan é dada por Acemoglu (2008, cap. 2). O Apêndice A é uma versão resumida do modelo descrito nessa obra. Por conveniência, reproduz-se aqui a condição de estado estável do modelo com crescimento demográfico e a equação diferencial de equilíbrio, em tempo discreto:

$$f[k^*(t)]/k^*(t) = n + \delta/s, \quad (2.1)$$

e

$$k(t+1) = sf[k(t)] - (n + \delta)k(t). \quad (2.2)$$

onde  $k$  é o estoque de capital por trabalhador,  $n$  a taxa de crescimento demográfico,  $\delta$  a taxa de depreciação do capital,  $s$  a taxa de poupança, e  $t$  o indicador de tempo.

Não é difícil de perceber que no estado estável de um modelo como este a renda per capita de um país depende somente de fatores exógenos. Ademais, somente um destes fatores é ilimitado, possibilitando, assim, o crescimento contínuo da renda per capita: a tecnologia. Observe que a tecnologia aqui incorpora todos os fatores que entram indiretamente na função de produção, não apenas o “como fazer”, mas também capacidades organizacionais, estabilidade política e uma justiça que funcione, entre outros. Este resultado é, como já foi notado, exatamente onde se encontra a fraqueza deste modelo ou do seu poder insuficiente de explicação. O resultado mais importante neste modelo, crucial para o desenvolvimento econômico, depende exclusivamente de variáveis exógenas. A conclusão não muda quando se inclui a otimização intertemporal de utilidade das famílias no modelo para fechar o modelo *neoclássico* de crescimento econômico.

Por que o modelo não permite crescimento econômico irrestrito sem contar com o progresso técnico exógeno? A resposta é tanto simples quanto desafiante. A parte simples é: A causa são os retornos decrescentes do capital e do trabalho. Isto nos conduz à parte complexa: Como surgem os retornos constantes ou mesmo crescentes de algum fator? Onde está a falha na lei de rendimentos decrescentes?

Uma maneira de permitir crescimento irrestrito no modelo Solow-Swan especificado na seção precedente é alterar um de seus pressupostos: Os retornos do capital devem ser constantes em vez de serem decrescentes. Esta mudança permite o crescimento prolongado da renda per capita mesmo no modelo neoclássico, mas viola as condições de Inada e deixa-nos sem uma explicação do progresso tecnológico. Negligenciaremos este caso.

A outra maneira é reintroduzir o progresso técnico no modelo. Como isto é feito, entretanto, é uma questão importante. Como mostra Acemoglu (2008, seção 2.7), dependendo da maneira que se introduz  $A(t)$  na função de produção geral, o crescimento *equilibrado* não pode ser garantido. O crescimento equilibrado faz referência a uma alocação em que o produto cresce a uma taxa constante e a taxa de juro e as proporções de fatores per-

manecem constantes. É possível mostrar que somente um tipo de progresso tecnológico – isto é, Harrod-neutro (puramente ampliador de trabalho) – pode ser conciliado com uma trajetória de crescimento equilibrado (do produto, do capital, e do consumo) e com uma função de produção com retornos constantes de escala para  $K$  e  $L$ . Em outras palavras, é preciso que possamos representar a função de produção através da forma

$$Y(t) = F[K(t), A(t)L(t)], \quad (2.3)$$

se quisermos garantir que exista uma trajetória de crescimento equilibrado após algum período  $T$ . A explicação intuitiva para este resultado vem do pressuposto da acumulação de capital: Numa trajetória de crescimento equilibrado, tanto o capital quanto o produto devem crescer à mesma taxa; para existir progresso tecnológico e crescimento de renda per capita, essa taxa de crescimento deve ser maior do que a taxa de aumento da população; isto conduz a uma relação capital-trabalho crescente e, assim, a uma assimetria entre capital e trabalho; por conseguinte, a constância de crescimento requer que a mudança tecnológica tome uma forma expansiva de trabalho (*labor augmenting*).<sup>1</sup>

Um passo ulterior para melhorar significativamente o “ajuste” do modelo de Solow aos dados reais é introduzir o capital humano nele [ver, por exemplo, Mankiw, Romer e Weil (1992)]. O *capital humano* pode ser entendido como resultante de todo investimento feito por pessoas que eleve a sua produtividade. Exemplos de investimentos que tendem a elevar a produtividade do trabalho são educação formal e informal, treinamento no local de trabalho, melhoramento da saúde e migração para obter oportunidades de trabalho melhores.

Ainda que a introdução de capital humano no modelo de Solow melhore significativamente o seu ajuste aos dados, as diferenças tecnológicas ainda explicam grandes diferenças entre os países quanto à renda per capita. Ademais, é razoável supor que o progresso tecnológico está associado fortemente a investimentos em capital humano e atividades de pesquisa, e fundamentalmente a instituições apropriadas – um pressuposto que é irreconciliável com o progresso tecnológico ortogonal. O modelo de Solow não é capaz de dar uma resposta a estas questões. E esta é a principal razão para a emergência dos modelos de crescimento *endógenos*.

---

<sup>1</sup>Também é interessante observar que em modelos de crescimento endógenos a mesma intuição implica que a tecnologia deve ser endogenamente mais ampliadora de trabalho do que de capital após algum instante  $T$  [veja Acemoglu (2008, cap. 15)].

## 2.2 Modelos de crescimento endógeno

Os modelos de crescimento endógenos têm a sua origem em meados dos anos 1980, embora certamente houve trabalhos precedentes. A contribuição principal veio de Romer (1990), mas também de Romer (1986) e Lucas (1988). A ideia fundamental por trás destes modelos é a necessidade de explicar o processo de mudança tecnológica a fim de entender diferentes tendências históricas de crescimento da renda per capita entre países. Os modelos precedentes de crescimento “exógeno” foram capazes de explicar algumas tendências interessantes do crescimento de longo prazo dos países, mas se baseou demasiadamente em uma fonte exógena de crescimento, isto é, a tecnologia num sentido amplo. Nesta seção se vê por que a endogeneização do progresso tecnológico constituiu um passo difícil mas essencial para a construção de modelos de crescimento, e como a dificuldade está diretamente relacionada a aspectos intrínsecos do conhecimento e de sua difusão.

Como notado antes, a tecnologia no seu senso estrito é composta de conhecimento, um bem não-rival e pelo menos quase não-excluível. Por isto, ele é intrinsecamente diferente do capital (seja este humano ou físico) ou do trabalho. Num primeiro olhar, o capital humano pode parecer ser também uma forma de conhecimento, mas uma perspectiva cuidadosa revela que ele é inseparável dos seres humanos e, portanto, um bem rival e excluível.

A maneira mais simples de tornar a tecnologia endógena é fazendo-a uma função do tamanho da população: A existência de um maior número de pessoas aumenta a probabilidade de novas idéias surgirem e, dessa forma, acelera o progresso tecnológico. Mas a observação da realidade torna imediata a percepção de que há muitos exemplos de como esta lógica não se aplica. A introspecção fundamental vem da economia de ideias, formalizado por Romer (1986, 1990).

Supusemos até agora que a tecnologia fosse um bem não-rival e completamente não-excluível. A ideia introduzida por Romer é que o conhecimento (“ideias”) é não-rival e pelo menos parcialmente excluível, resultando largamente do comportamento intencional de agentes econômicos que buscam o lucro. O exemplo de um insumo tecnológico não-rival mas excluível, citado por Romer, é o projeto de um novo bem de capital. Embora a vasta maioria desses projetos resulte das atividades de pesquisa e desenvolvimento de empresas privadas maximizadoras de lucro, eles são todos não-rivais porque eles podem

ser facilmente reproduzidos e usados por muitas empresas e indivíduos ao mesmo tempo.

A característica de não-rivalidade de ideias é importante porque implica um custo fixo de produção e um custo marginal nulo. Uma vez que uma nova ideia foi gerada, a sua reprodução não tem nenhum custo além do meio físico, que normalmente é um bem rival. Um custo fixo de produção e um custo marginal baixo e constante implicam um custo médio decrescente e retornos crescentes de escala. Já que o custo médio é sempre maior do que o de custo marginal, a precificação com base no custo marginal leva a lucros negativos. Como efeito, nenhuma companhia investiria recursos numa nova ideia se não pudesse obter nenhum poder de fixação de preços com a invenção. Assim, a economia de ideias naturalmente conduz ao abandono da competição perfeita. (JONES, 2000).

Antes de Romer, outros salientaram a importância de externalidades de capital humano [por exemplo, Arrow (1962), Lucas (1988)] e do setor de pesquisa [por exemplo, Uzawa (1965)] no processo de crescimento econômico. A diferença é que todos eles trataram o conhecimento como um bem não-rival e *totalmente* não-excluível, gratuita e imediatamente disponível para todos os agentes depois da sua geração. Há boas razões, entretanto, para o pressuposto de exclusibilidade parcial do conhecimento e da tecnologia. Não apenas a origem da legislação de patentes coincide com a origem da rápida aceleração de crescimento que se observa nos últimos dois séculos, mas a exclusibilidade parcial também explica por que muitas companhias modernas estão dispostas a investir ativamente em pesquisa e inovação. Se não houvesse uma possibilidade de exclusão, as empresas não investiriam numa nova tecnologia que poderia ser se apropriada imediata e gratuitamente por outros. O obstáculo concreto seria a incorrência em um custo fixo para obter um lucro nulo.

De acordo com Romer (1990, p. S75), as implicações da não-rivalidade para a teoria de crescimento econômico são duas: “First, nonrival goods can be accumulated without bound on a per capita basis. [...] Second, treating knowledge as a nonrival good makes it possible to talk sensibly about knowledge spillovers, that is, incomplete excludability.”.

Esses aspectos estão relacionados intrinsecamente com a convexidade da função de produção. A existência de um insumo não-rival conduz a uma função de produção com retornos crescentes de escala, isto é, convexa em relação à origem. Isto por sua vez viola o teorema de Euler e torna impossível que todos os insumos sejam compensados pela sua contribuição marginal ao produto. Há pelo menos cinco maneiras para resolver isto: pode-

se supor que a tecnologia esteja exógena e livremente disponível (SOLOW, 1956); que a tecnologia seja provida por um governo que recolhe tributos (SHELL, 1966); que a tecnologia resulte da alocação de recursos entre um setor de pesquisa e um setor de bens finais (UZAWA, 1965); que a tecnologia seja um subproduto não-intencional da acumulação de outros fatores (capital, por exemplo humano) (ARROW, 1962; LUCAS, 1988); ou que a tecnologia seja inventada e aperfeiçoada pelo comportamento maximizador de firmas privadas que buscam obter rendas de monopólio (ROMER, 1990). Enquanto todos os procedimentos tornam o insumo não-rival tecnologia essencial ao crescimento econômico, e alguns deles (Arrow, Uzawa, e Lucas) enfatizam a importância das externalidades do capital humano, somente a estratégia de Romer oferece uma explicação baseada no mercado dos investimentos em progresso tecnológico. Por causa desta vantagem, e por causa do interesse específico do presente estudo, detalha-se no Apêndice B a estratégia de modelagem de Romer.

Um ponto forte do modelo de Romer (1990) é a tentativa de fundamentação microeconômica de um modelo de crescimento econômico que torna o progresso tecnológico endógeno. O modelo de Romer torna o progresso tecnológico endógeno fazendo-o resultado da busca de novas idéias por pesquisadores interessados em tirar proveito econômico delas. Especificamente, o modelo de Romer é sobre adiantar a fronteira tecnológica, isto é, sobre o progresso tecnológico conduzido por pesquisa e desenvolvimento no mundo avançado.

A característica crucial de como Romer especifica o nível tecnológico é que o conhecimento afeta a produção de duas formas. Um novo projeto permite a produção de um novo bem de capital, o qual pode ser usado para gerar produto. Um novo projeto também eleva o estoque de conhecimento e, conseqüentemente, a produtividade do capital humano no setor de pesquisa. O proprietário de um projeto tem direitos de propriedade sobre o seu uso na produção de bens de capital mas não sobre o seu uso em pesquisa. Neste sentido, o conhecimento é parcialmente excluível.

Embora a origem dos retornos não-decrescentes seja o efeito de espraiamento do conhecimento, o nível tecnológico,  $A$ , afeta somente indiretamente o setor de bens finais por causa dos seus efeitos sobre a disponibilidade de novos bens de capital. No setor de pesquisa,  $A$  entra diretamente.

Como neste modelo uma fração constante da população é empregada na geração de

ideias, a taxa de crescimento da renda per capita ao longo de uma trajetória de crescimento equilibrado é igual à taxa de crescimento da proporção capital-trabalho e igual à taxa de crescimento do estoque de ideias, como no modelo de Solow. A diferença em relação ao último é que no modelo de Romer a taxa de crescimento do estoque de conhecimento é endógena e dada por

$$g_A = \delta H_A. \quad (2.4)$$

Romer supõe que a produtividade da pesquisa seja proporcional ao estoque existente de ideias:  $\bar{\delta} = \delta A$ . Com este pressuposto, a produtividade dos pesquisadores cresce no tempo ainda que o número de pesquisadores seja constante.

Apesar de sua atratividade teórica, a evidência empírica não tem apoiado essa hipótese; é mais provável que  $\delta$  seja uma função côncava de  $A$ , como arguido por Jones (1995). Com essa nova hipótese, o modelo de Romer implica que políticas de aumento do nível de investimento ou de aumento da proporção de pesquisadores, por exemplo, não impactam no crescimento econômico de longo prazo, exatamente como no modelo neoclássico. Estas políticas teriam então um impacto somente sobre a trajetória de crescimento (efeito de nível) entre um estado estável e outro, e não sobre a taxa de crescimento no novo estado estacionário (efeito de crescimento).

Mesmo assim continua havendo uma diferença importante. Ao contrário do modelo de Solow, no modelo de Romer nem todos os insumos recebem o equivalente ao seu produto marginal; uma situação que é possível apenas em uma economia em competição imperfeita. No caso do modelo de Romer, as companhias no setor de bens intermediários recebem um preço superior ao custo marginal, passando, contudo, o excedente aos inventores.

## 2.3 Externalidades nos modelos de crescimento

Se aceito que a função de produção seja homogênea de grau um a respeito dos insumos limitados (capital e trabalho), a única fonte potencial de crescimento de longo prazo é um insumo tecnológico ilimitado, o qual deve adentrar a função de produção com retornos não-decrescentes. O principal problema na modelagem da mudança tecnológica endógena é, conseqüentemente, especificar de onde os retornos não-decrescentes da tecnologia (ou do conhecimento) provêm.

Como notado na seção precedente, existem basicamente duas maneiras pelas quais

os retornos não-decrescentes da tecnologia podem surgir. Enquanto a abordagem de incentivos de mercado é muito mais sensata, a abordagem de externalidades é ainda assim complementar porque em ambos os casos as externalidades surgem do caráter não-rival do conhecimento. A abordagem de incentivos de mercado requer somente que o conhecimento seja *parcialmente* excluível.

No modelo de Romer (1990) os retornos não-decrescentes da tecnologia vêm explicitamente da equação de acumulação de conhecimento,

$$\dot{A} = \delta H_A A, \quad (\text{B.3})$$

que é uma função linear do capital humano destinado à pesquisa e, mais importante, do nível tecnológico já alcançado. A explicação é que o conhecimento acumulado torna os pesquisadores de hoje mais produtivos que os de, digamos, cinquenta anos atrás. Isto significa de fato que há efeitos de espraiamento do conhecimento no setor de pesquisa da economia, como enfatizado por Romer (1990, p. S83) quando ele afirma que “All researchers can take advantage of  $A$  at the same time”.

Em contraste, no modelo de Lucas (1988) os retornos não-decrescentes vêm da equação de acumulação de capital humano

$$\dot{h} = h(t)^\zeta G[1 - u(t)], \quad (\text{2.5})$$

que simplifica para

$$\dot{h} = h(t)\delta[1 - u(t)], \quad (\text{2.6})$$

Onde  $1 - u(t)$  é o tempo dedicado à acumulação de capital humano,  $\zeta$  é supostamente igual a um e coerente com a evidência providenciada por Rosen (1976), e se supõe que a função  $G$  seja linear por simplicidade. O capital humano adentra a função de produção através de um insumo de trabalho efetivo  $e$  de um termo de capital humano médio que adentra a função de produção porque se supõe que a produtividade individual seja influenciada positivamente pelo capital humano médio. A ideia subjacente é que trabalhadores muito produtivos (com elevado capital humano) afetem positivamente a produtividade de trabalhadores menos produtivos (com baixo capital humano) que estão a sua volta.

O principal problema com o modelo de Lucas é que ele assume implicitamente um estoque ilimitado de capital humano, algo que é improvável se entendermos o capital humano como definido por Mankiw (1995, p. 298):

Knowledge refers to society's understanding about how the world works. Human capital refers to the resources expended transmitting this understanding to the labor force. Put crudely, knowledge is the quality of society's textbooks; human capital is the amount of time that has been spent reading them.

Outra maneira de formular a acumulação de capital humano é dada por McDermott (2002, p. 8), que o faz uma função do “grau de especialização por unidade de trabalho”,  $M/e_M n$ , do nível de capital humano,  $h$ , e da fração de esforço dedicado à pesquisa,  $e_L$ :

$$\dot{h} = \left( \frac{M}{e_M n} \right)^\gamma h^{1-\gamma} e_L. \quad (2.7)$$

O parâmetro  $\gamma$  é menor que um e o “grau de especialização” é dado por

$$M = b_1 e_M a h_a n N_c, \quad (2.8)$$

onde  $e_M$  é a quantia de esforço dedicado a trabalhar no setor de “mercado” (em oposição ao setor “primitivo”),  $h_a$  é o estoque médio de capital humano individual,  $n$  é o número de famílias que povoam cada “cidade”, e  $N_c$  é o número de “cidades” na economia.

Em outro artigo, McDermott (1999) especifica a acumulação de capital humano como uma função do seu estoque individual,  $h$ , da fração do esforço diário destinado à aprendizagem,  $e_L$ , e da “tecnologia mundial”,  $\Omega$ , a qual cresce exogenamente:

$$\dot{h} = h^{1-\gamma} (\kappa \Omega)^\gamma e_L. \quad (2.9)$$

Este tipo especificação clareia o papel que o capital humano desempenha na absorção de conhecimento pelas pessoas. Quanto mais elevado for o estoque de capital humano e quanto maior for o esforço diário de aprendizagem, maior é a capacidade do indivíduo absorver o conhecimento global e gratuitamente disponível e de aumentar o seu próprio capital humano.

A mesma idéia está presente no modelo desenvolvido por Easterly et al. (1994), em que os trabalhadores mais (ou melhor) qualificados podem usar mais (e mais avançados) bens de capital que trabalhadores não qualificados, e a acumulação de capital humano é dada pela equação

$$\dot{h} = \mu e^{\varphi u} A^\gamma h^{1-\gamma}, \quad (2.10)$$

em que  $u$  indica o tempo que uma pessoa se dedica à acumulação de habilidades em vez de trabalhar,  $A$  representa a fronteira tecnológica global, e se supõe que os parâmetros

atendam às condições  $\mu > 0$  e  $0 < \gamma \leq 1$ . Novamente, a “absorção” de conhecimento global por qualquer país depende do seu esforço dedicado à acumulação de capital humano e do seu verdadeiro nível de capacidades.

A vantagem da abordagem de *absorção de tecnologia* surge do fato dela permitir taxas positivas de crescimento de longo prazo da renda per capita em países específicos sem requerer retornos constantes para qualquer insumo de produção. Além disso, adapta-se melhor aos países não avançados que empurram um tanto fracamente a fronteira tecnológica global, mas que ainda podem tirar proveito do conhecimento criado em outro lugar, dedicando esforço à aprendizagem e à aquisição de habilidades. O modelo implica, ademais, que os países em via de desenvolvimento podem melhorar significativamente a sua riqueza dedicando simplesmente mais recursos ao aperfeiçoamento de capacidades genéricas (ensino primário e secundário) e específicas (ensino superior; treinamento profissional), sem a necessidade de investir demasiadamente em pesquisa, a qual frequentemente é muito cara para esses países. Uma observação empírica que este modelo não pode explicar são as diferenças de produtividade total de fatores que se observam entre países.

Enquanto em Romer (1990) o conhecimento tem retornos crescentes, mas com a sua taxa de crescimento limitada pelos retornos decrescentes da atividade de pesquisa, Lucas (1988) introduziu em seu modelo uma ideia presente na literatura econômica desde Marshall (1890) no mais tardar: o efeito externo do capital humano. Lucas argumenta que uma porção do capital humano médio da população é absorvida pelo indivíduo através da interação social, e isto seria o efeito externo do capital humano, isto é, uma externalidade positiva associada à acumulação de capital humano. A soma de ambos os efeitos do investimento em capital humano, o interno e o externo, teria um efeito positivo sobre a produção agregada.

A respeito do âmbito da interação social requerida para dar origem a efeitos externos de capital humano, Lucas cita o trabalho de Jacobs (1969) para argumentar a favor das cidades como foco de análise. O que Jacobs sugere é que a concentração da atividade econômica em cidades foi em parte o resultado das externalidades de capital humano derivadas da troca de ideias entre trabalhadores e empresários. Contudo, se este for o caso, então é mais provável que essas externalidades surjam dentro de limites setoriais do que geográficos, porque é esse tipo de interação que permite falar sensatamente sobre

“aprender-interagindo” – pelo menos depois da revolução dos meios de comunicação e de transporte que o mundo viu no último século.

De fato, as externalidades em Romer (1990) podem ser percebidas como se originando da interação humana num setor específico: o setor de pesquisa. No seu modelo simplificado, as externalidades da acumulação de conhecimento devem surgir lá porque é onde ocorre o processo de aprendizagem. Mas, como exposto na Introdução, a ideia é que as externalidades se originem onde quer que aconteça a aquisição de conhecimento, seja em escolas, universidades, instituições de pesquisa, departamentos de pesquisa e desenvolvimento dentro de empresas ou em outras atividades econômicas intensivas em capital humano.

A seguir são apresentadas algumas tentativas de confirmar empiricamente a existência de externalidades tecnológicas do capital humano na forma como foram sugeridas por Jacobs (1969) e Lucas (1988), bem como esforços de medição para quantificar essas externalidades e o próprio estoque de capital humano. Subsequentemente, será retomada a possibilidade de haver externalidades setoriais de capital humano e efeitos de espraiamento do conhecimento a nível setorial, o principal foco do presente estudo.

## 3 Evidências empíricas

### 3.1 Externalidades do capital humano

Uma das primeiras tentativas de mensurar os efeitos externos tecnológicos de capital humano foi feita por Rauch (1993), que estimou regressões de rendimento quase-mincerianas. A diferença principal em relação à equação minceriana é a inclusão do capital humano médio dos trabalhadores no mercado de trabalho local no lado direito da equação. Especificamente, Rauch estimou modelos da forma:

$$\ln W_{j,m} = X'_{j,m}\beta + \gamma_p S_{j,m} + \gamma_e S_m, \quad (3.1)$$

onde  $X_{j,m}$  é um vetor de variáveis de controle,  $S_{j,m}$  são os anos de escolaridade do indivíduo  $j$  que vive/trabalha no mercado de trabalho  $m$ ,  $s_m$  é a média de anos de escolaridade dos trabalhadores no mercado de trabalho  $m$ ,  $\gamma_p$  mede o *retorno privado* da educação e o coeficiente  $\gamma_e$  mede o *retorno externo* da educação.

Os resultados de Rauch mostram elevados retornos externos à instrução formal, frequentemente mais elevados que os retornos privados. Entretanto, como notado por Acemoglu (2008), as regressões de Rauch utilizam diferenciais médios de escolaridade entre cidades, os quais podem refletir muitos fatores que também afetam diretamente os salários (um alto custo de vida, por exemplo). Estudos mais recentes, tais como Acemoglu e Angrist (2001), Duflo (2004) e Ciccone e Peri (2006), estimam que os efeitos externos do capital humano sejam relativamente pequenos (até mesmo insignificantes) em mercados de trabalho locais (ACEMOGLU, 2008). Moretti (2004) encontrou externalidades mais elevadas.

Embora a evidência empírica sugira que os efeitos externos do capital humano sejam pequenos, os estudos feitos até agora abordaram principalmente os efeitos sobre mercados de trabalho locais, tal como cidades ou, no caso de Acemoglu e Angrist (2001), estados. Quase não há estudos que calcularam estes efeitos para níveis mais agregados, tais como setores, países e grupos de países.

Um dos poucos estudos que calculou o retorno externo à escolaridade por setores

econômicos é o de Heuermann (2009). Além de encontrar retornos mais altos para a indústria do que para os serviços – algo que o autor relaciona às externalidades pecuniárias decorrentes da intensidade de capital físico na indústria – Heuermann estimou retornos de 1,8% para um aumento de um ponto percentual na proporção de trabalhadores altamente qualificados e de 0,6% para um aumento equivalente na proporção de trabalhadores que não são altamente qualificados. Outro resultado interessante é que os trabalhadores altamente qualificados tendem a tirar proveito de efeitos de espriamento intra-setoriais do capital humano, enquanto os trabalhadores não altamente qualificados tendem a tirar proveito principalmente das externalidades *pecuniárias*.

A potencial natureza setorial das externalidades do capital humano é enfatizada por Wheaton e Lewis (2002), que encontrou que graus mais altos de especialização e concentração setorial estão associados a ganhos salariais, sinalizando que as externalidades do capital humano podem ser limitadas pelo âmbito setorial.

Para isolar efeitos potencialmente distintos, Wheeler (2007) estima equações de salário contendo tanto uma medida do nível de capital humano agregado ou médio quanto uma medida da concentração espacial dos setores, e achou que ambos os efeitos são elevados e significantes. Este resultado lhe fez concluir que economias locais e externalidades de capital humano parecem ser fenômenos bastante distintos.

Esta descoberta sugere que nem as externalidades de capital humano tecnológicas nem as pecuniárias se originem somente com a concentração espacial. Uma extensão lógica desta análise é a investigação de trabalhadores que mudam de setor, mas dentro da mesma cidade. A evidência providenciada por Jaffe, Trajtenberg e Henderson (1993) de que o efeito de difusão geográfica do conhecimento é limitado também sustenta esta abordagem. Assim como o estudo de Rosenthal e Strange (2008), que confirmou que a densidade de trabalhadores altamente educados está associada fortemente aos níveis de salário, e que esta associação é reduzida por um fator de 2,5 a 3,0 quando se excede um círculo concêntrico de cinco milhas (8 km) em torno do lugar de trabalho do indivíduo. Fu (2007) relata um enfraquecimento muito rápido dos efeitos do capital humano até mesmo além de um círculo de três milhas (4.8 km).

Charlot e Duranton (2004), por sua vez, chama atenção para a importância dos meios de comunicação em mediar os efeitos de aglomeração humana e concentração do capital humano sobre os salários. De acordo com os autores, controlar para a intensidade da

comunicação no lugar de trabalho reduz os efeitos diretos em oito a dez vezes. Esta evidência também sugere que os efeitos de espraiamento do capital humano são mais fortes dentro de setores do que entre setores ou em uma área geográfica específica.

Finalmente, é possível que para perceber as externalidades espaciais de capital humano seja necessário considerar o valor de consumo que as áreas com elevado capital humano têm para as pessoas. Neste caso, é recomendado analisar o comportamento dos aluguéis também, além dos salários, como faz Rauch (1993), por exemplo. A seguir são apresentadas algumas evidências que relacionam as externalidades do capital humano com o retorno salarial da educação.

## 3.2 Retornos da escolarização

Outra indicação da existência de efeitos externos associados à acumulação de capital humano são os resultados encontrados por Card e Krueger (1992), Heckman, Layne-Fanar e Todd (1996), Jaeger e Page (1996) e Dias e McDermott (2003), entre outros. Estes autores perceberam que a relação entre anos de escolaridade e renda individual não é linear. Existe um “efeito pergaminho” (ou efeito diploma), que aumenta drasticamente o retorno salarial de indivíduos que completam um dado título acadêmico, e um “efeito limiar”, que muda a declividade da relação entre escolaridade e renda quando um certo nível de escolaridade é alcançado.

O efeito diploma ocorre provavelmente devido a um mecanismo de sinalização da produtividade, e é frequentemente superestimado (ANTELIUS, 2000; FLORES-LAGUNES; LIGHT, 2004). Além do mais, este efeito se reduz à medida que o empregador obtém informações adicionais sobre a produtividade dos seus empregados [veja Belman e Heywood (1997)].

O efeito limiar é mais importante porque, como indicado por Dias e McDermott (2003), parece que as externalidades positivas da escolarização sobre a economia se originam somente após a acumulação de um certo nível mínimo de estoque de capital humano. Aplicando um modelo macro-minceriano a dados da Pesquisa Mundial de Valor (*World Value Survey*) que abarcam o período 1960-1990, os autores estimaram um valor limiar de 4,5 anos de escolaridade para um conjunto de 109 países. Para o Brasil, o valor estimado foi de 3 anos. Ressalta-se que este valor limiar se refere ao nível educacional médio ou

agregado da economia, e não ao nível individual.

Para o estado brasileiro de Minas Gerais, Hoffmann e Simão (2005), usando dados do Censo de 2000, e Salvato e Silva (2008), usando dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2005, encontraram um valor limiar de 10 anos de escolaridade após o qual anos adicionais de estudo elevam drasticamente o rendimento individual. A relação não linear entre renda e escolaridade no Brasil foi observada também por Lam e Schoeni (1993) e Ramos e Vieira (1996), entre outros.

Em uma abordagem diferente, Trostel (2004) modifica a equação de Mincer (1974) para permitir uma relação em U invertido entre escolaridade e renda. Ao mesmo tempo, a especificação de Trostel permite verificar o retorno de escala da educação com uma perda mínima de graus de liberdade<sup>1</sup>. Implementando essa especificação aos estados do Brasil, com dados da PNAD de 2004 a 2008, Monteiro, Dias e Dias (2009) confirmaram a presença de retornos crescentes da educação após um mínimo de 4 anos de educação.

As evidências sobre retornos da escolarização por setores econômicos são claramente insatisfatórias. No Brasil, Hoffmann e Simão (2005) estimam os retornos da educação para três grandes setores, utilizando dados do Censo 2000 para o estado de Minas Gerais, e encontram que: (i) abaixo de 10 anos de escolaridade, a taxa de retorno para cada ano adicional de escolaridade é de 5,4% na agricultura, de 6,9% na indústria e de 8,7% nos serviços; (ii) para indivíduos com pelo menos 10 anos de escolaridade, as taxas de retorno são, respectivamente, 19,8%, 23,3% e 21,4%. Deve ser notado que, como na agricultura a proporção de pessoas com escolaridade acima de 10 anos é de somente 4,7% (contra 19,6% na indústria e 37,1% nos serviços), a variação da renda devida à instrução é muito limitada neste setor.

---

<sup>1</sup>Outra opção, bastante comum, é estimar coeficientes para cada nível ou mesmo para cada ano de escolaridade.

## 4 Mensuração do capital humano

Ao desenvolver a estrutura teórica do que seria denominado “teoria do capital humano”, Schultz (1961) e Becker (1962) estavam tentando relacionar a teoria de crescimento neoclássica com evidências empíricas que apontavam para um crescimento da renda per capita muito mais elevado do que a taxa de crescimento do capital físico. Eles estavam preocupados em superar a barreira moral contemporânea de enxergar seres humanos como um insumo produtivo e em mostrar por que a aquisição de habilidades e capacidades de trabalho pelas pessoas deveria ser considerada um investimento produtivo que levaria à acumulação de um tipo de capital, o capital humano. Providenciando uma lógica microeconômica para o investimento de uma pessoa em sua capacidade produtiva, eles deslocaram a atenção do investimento *em* pessoas, que implicava posse e era rememorativa de escravidão, para o investimento *pela* pessoa *nela mesma*, que implicava capacidade de escolha individual.

Como observou Schultz (1961), muito do que parece o consumo de um indivíduo é de fato um tipo de investimento na sua capacidade de obter renda. Entre os exemplos citados por Schultz estão a instrução formal, programas de estudo para adultos, cuidados de saúde, migração interna para aproveitar oportunidades de trabalho melhores, rendimentos aos quais estudantes adultos renunciam para poder continuar frequentando a escola, e treinamentos no local de trabalho. Enquanto a distinção entre despesas para consumo e para investimento é muito difícil na prática, ele defende a medição do investimento humano pelo seu produto em vez de pelo seu custo:

While any capability produced by human investment becomes a part of the human agent and hence cannot be sold; it is nevertheless 'in touch with the market place' by affecting the wages and salaries the human agent can earn. The resulting increase in earnings is the yield on the investment. In principle, the value of the investment can be determined by discounting the additional future earnings it yields just as the value of a physical capital good can be determined by discounting its income stream. (SCHULTZ, 1961, p. 8).

Para Schultz, a educação poderia não apenas ser medida mais facilmente (anos de instrução) que outros componentes do capital humano, mas seria também o fator mais importante para a acumulação de capital humano. Contudo, ele aconselhou isto porque

todos os outros campos necessários da economia estavam ainda na sua infância, algo que não se aplica mais.

Uma vez que o capital humano foi estabelecido como capital, o passo seguinte era encontrar uma maneira de ligar investimento em capital humano a ganhos reais na produtividade. Como demonstrado por Mincer (1958, 1962, 1974) e Becker (1962), é possível estabelecer uma relação entre investimento em capital humano e ganhos de produtividade observando o comportamento do padrão de renda de um indivíduo ao longo de seu ciclo de vida e o nível educacional que alcançou, pelo menos sob determinadas condições. Entre os pressupostos mais importantes está a existência de um mercado de trabalho competitivo, uma influência relativamente pequena de diferenças de capacidade inatas, que a educação seja um dos principais investimentos em capital humano e que este investimento seja exógeno, isto é, que não seja dependente da renda.

Até agora foi simplesmente ignorada a questão de se o nível de instrução individual é uma aproximação boa para o estoque individual de capital humano. Embora medir o capital humano usando anos de escolaridade seja uma abordagem muito comum, deve-se ter consciência de suas limitações, como foi salientado recentemente por muitos pesquisadores. A principal complicação surge dos aspectos qualitativos da instrução. Como mostrado primordialmente por Behrman e Birdsall (1983), e por outros depois, ignorar os efeitos da qualidade da instrução sobre a acumulação de capital humano pode viesar drasticamente as estimativas do retorno da escolaridade por meio de equações de rendimento de tipo minceriano. Os autores propõem uma equação minceriana adaptada para a qualidade da educação, a qual poderia, pelo menos para o caso brasileiro, ser aproximada pela educação média de professores de escolas primárias e secundárias na cidade ou estado onde estudou o indivíduo [para uma discussão, veja Heyneman e Loxley (1983)]. Mais recentemente, Hanushek (2009) declarou que as capacidades cognitivas, medidas por provas de desempenho, têm um impacto claro e forte sobre o rendimento pessoal e sobre o produto agregado de países pela mudança em sua taxa de crescimento. De acordo com Hanushek e Kimko (2000), adicionando uma medida da qualidade educacional (desempenho na prova de matemática PISA) à especificação básica que inclui somente a renda inicial e a educação média eleva a proporção da variabilidade do PIB per capita dos 31 países da amostra que podem ser explicados pelo modelo de 33% para 73%. Adicionalmente, a introdução da medida de qualidade educacional no modelo reduz drasticamente o efeito dos anos de escolaridade sobre a renda per capita, e o torna insignificante para a

---

maioria dos países.

Quanto aos insumos da educação, Hanushek (2010) argumenta que o componente que mais afeta os resultados do processo educacional é a qualidade dos professores. Porém, o autor também mostra que não há nenhuma prova de que a qualidade esteja relacionada de forma consistente com características observáveis geralmente disponíveis. Ignoramos aqui esta complicação, sobretudo porque o desenho amostral dos nossos dados possui representatividade somente em nível estadual, e não seria razoável assumir que a qualidade da educação seja homogênea a esse nível.

## 5 Procedimentos metodológicos

### 5.1 Implementação empírica

Em vez de derivar um modelo empiricamente testável diretamente de qualquer teoria mencionada, o modelo a seguir é derivado a partir de uma equação salarial genérica caracterizada pelo equilíbrio não-walrasiano de Acemoglu (1996), acrescida de variáveis que captem efeitos de transbordamento do conhecimento e cuja inspiração provém dos trabalhos de Romer (1990) sobre externalidades setoriais e de Lucas (1988) sobre externalidades do capital humano.

Suponha-se uma equação de salário sujeita ao equilíbrio não-walrasiano em que o salário seja uma proporção do produto,

$$w_{ij} = \rho f[H_i(S, X), A_j(R_j), Z], \quad (5.1)$$

que descreva o salário real do indivíduo  $i$ , empregado no setor  $j$ , como uma função de seu próprio capital humano,  $H_i$ , o nível de conhecimento no setor onde ele é empregado,  $A_j$ , e um vetor de variáveis que descrevem o mercado de trabalho,  $Z$ . O nível individual de capital humano é uma função da escolaridade do indivíduo,  $S$ , e de sua experiência no mercado de trabalho,  $X$ . O nível de conhecimento no setor  $j$  é uma função da proporção de pessoas envolvidas em atividades de pesquisa nesse setor,  $R_j$ . As derivadas parciais de primeira ordem  $\partial f/\partial H_i$  e  $\partial f/\partial A_j$  são ambas positivas.

O crescimento do conhecimento setorial ocorre como exposto em Romer (1990):  $g_A = \delta R_j$ . O nível de conhecimento em qualquer setor no tempo  $t$  é dado pela equação exponencial  $A_j = A_0 e^{g_A t}$ , onde  $A_0$  é o nível de conhecimento inicial. O capital humano individual é igual a  $H_i = \gamma e^{\Omega(S) + \Phi(X)}$ .

Assumindo a forma funcional

$$f[H_i(S, X), A_j(R_j), Z] = A_j^\alpha H_i^\beta e^{\chi Z} \quad (5.2)$$

para a função de produção, tem-se que o salário do indivíduo  $i$ , empregado no setor  $j$ , é

$$w_{ij} = \rho A_j^\alpha H_i^\beta e^{\chi Z}. \quad (5.3)$$

A partir desta equação de salário é possível chegar a uma equação de (log-)rendimentos estimável da forma

$$\ln w_{ij} = \gamma_0 + \gamma_1 S + \gamma_2 X + \gamma_3 Z + \gamma_4 R_j. \quad (5.4)$$

O procedimento é descrito a seguir e serve principalmente para deixar claras as hipóteses assumidas, em alguns casos bastante restritivas. Assuma:

$$w_{ij} = f[H_i(S, X), A_j(R_j), Z] \quad \text{com} \quad \frac{\partial f}{\partial H_i} > 0 \quad \frac{\partial f}{\partial A_j} > 0 \quad (5.5a)$$

$$= \rho A_j^\alpha H_i^\beta e^{\chi Z} \quad (5.5b)$$

$$A_j = A_0 e^{g_A t} \quad (5.6)$$

$$g_A = \delta R_j \quad (5.7)$$

$$H_i = \gamma e^{\Omega(S) + \Phi(X)} \quad (5.8)$$

Substituindo (5.7) em (5.6), aplicando logaritmo e suprimindo o indicador de tempo:

$$\ln A_j = \ln A_0 + \delta R_j \quad (5.9)$$

Aplicando logaritmo em (5.8) e (5.5b):

$$\ln H_i = \ln \gamma + \Omega(S) + \Phi(X) \quad (5.10)$$

$$\ln w_{ij} = \ln \rho + \alpha \ln A_j + \beta \ln H_i + \chi Z \quad (5.11)$$

Substituindo (5.9) e (5.10) em (5.11):

$$\begin{aligned} \ln w_{ij} &= \ln \rho + \alpha(\ln A_0 + \delta R_j) + \beta[\ln \gamma + \Omega(S) + \Phi(X)] + \chi Z \\ &= (\ln \rho + \alpha \ln A_0 + \beta \ln \gamma) + \alpha \delta R_j + \beta \Omega(S) + \beta \Phi(X) + \chi Z \\ &= a + bS + cX + dZ + eR_j \end{aligned} \quad (5.12)$$

Onde

$$a \equiv \ln \rho + \alpha \ln A_0 + \beta \ln \gamma$$

$$b \equiv \beta \Omega$$

$$c \equiv \beta \Phi$$

$$d \equiv \chi$$

$$e \equiv \alpha \delta$$

Para estimar a equação (5.4) definimos  $S$  como sendo um termo cúbico de anos de escolaridade,  $X$  como sendo um termo quadrado de décadas de experiência no mercado de trabalho, e  $Z$  como um vetor das variáveis indicativas tipo de trabalho (trabalhador formal, militar, funcionário público, trabalhador informal, autônomo, empregador), cor (preta/parda, outra), área rural (sim/não), pessoa de referência da família (sim/não), região (Sudeste, Norte, Nordeste, Sul, Centro-Oeste), sindicalizado (sim/não), cargo (chefe, gerente, outro), região metropolitana (sim/não), casado (sim/não), e da variável discreta número de crianças menores de 14 anos na família. Para aproximar  $R_j$  usa-se a fração (em log) de pessoas em cada setor que tem pelo menos um mestrado (a partir de agora “concentração setorial de conhecimento”, representada por  $SKC$ ), e a fração (em log) de pessoas em cada estado empregadas no setor de serviços às empresas (a partir de agora “intensidade de difusão do conhecimento”, representada por  $KSI$ ).

Dessa forma, a primeira equação a ser estimada é

$$\ln w_{ij} = \gamma_0 + \Gamma_1 S + \Gamma_2 X + \gamma_3 SKC + \gamma_4 KSI + \Gamma_5 Z + \epsilon, \quad (5.13)$$

onde o último termo,  $\epsilon$ , é um termo de erro. Estimam-se os parâmetros  $\gamma_0, \gamma_3, \gamma_4$  e os vetores de coeficientes  $\Gamma_1, \Gamma_2, \Gamma_5$ .

Um possível viés de não-seleção surge porque se observa apenas o salário das pessoas que decidiram trabalhar, mas não daquelas pessoas cujo salário de reserva se encontra acima do salário de mercado. Usa-se o procedimento de Heckman (1979) para corrigir esta possível fonte de viés e incluem-se as seguintes variáveis na equação de seleção: renda familiar, número de membros da família, número de crianças menores de 14 anos na família, anos de escolaridade, idade do indivíduo, região, cor/etnia, casado ou não, mulher com pelo menos uma criança menor de 14 anos de idade.<sup>1</sup>

## 5.2 Dados

Todos os dados são da Pesquisa Nacional por Amostra de Domicílios (PNAD), volume de 2008, que foi baseada em entrevistas pessoais realizadas entre outubro e dezembro de 2008.<sup>2</sup> Desde 2004, a cobertura geográfica da pesquisa representa todo o país. O levantamento da PNAD tem um desenho amostral complexo e, portanto, requer um tratamento

<sup>1</sup>A rotina do Stata encontra-se no Apêndice C.

<sup>2</sup>Os microdados podem ser baixados em <http://www.ibge.gov.br>.

Tabela 5.1: Estatísticas descritivas

Variável	N	Soma	Média	DP	Min	Max	Pergunta(s) da PNAD
Variáveis contínuas							
Salário individual	145,640	148,000,000	1,015	1,467	1	30,000	4718 9058
Anos de escolaridade	198,467	1,633,748	8.11	4.40	0	17	4803 6003 6007
Idade/10	198,467	749,637	3.72	1.13	2	6	8005
Experiência/10	145,640	324,164	2.23	1.23	0	5.5	8005 9892
Renda familiar	198,467	425,000,000	2,112	2,993	0	158,500	4722
Membros da família	198,467	720,078	3.58	1.47	1	16	4724
Crianças na família	198,467	158,945	0.789	1.020	0	9	8005
SKC	140,828	-840,570	-5.969	1.877	-9.545	-2.023	9907 4803 6003 6007
KSI	201,398	-829,891	-4.121	0.450	-5.324	-3.082	9907 UF
Variáveis categóricas							
Feminino	201,398	102,487	0.509	0.500	0	1	0302
Empregado formal	145,640	60,239	0.414	0.492	0	1	4706
Militar	145,640	453	0.003	0.056	0	1	4706
Func. público	145,640	12,345	0.085	0.279	0	1	4706
Empregado informal	145,640	33,714	0.231	0.422	0	1	4706
Autônomo	145,640	31,947	0.219	0.414	0	1	4706
Empregador	145,640	6,942	0.048	0.213	0	1	4706
Preto/pardo	201,398	108,394	0.538	0.499	0	1	0404
Área rural	201,398	23,795	0.118	0.323	0	1	4728
Pessoa de refer.	201,398	94,275	0.468	0.499	0	1	0402
Sudeste	201,398	61,278	0.304	0.460	0	1	UF
Norte	201,398	24,621	0.122	0.328	0	1	UF
Nordeste	201,398	61,747	0.307	0.461	0	1	UF
Sul	201,398	30,596	0.152	0.359	0	1	UF
Centro-Oeste	201,398	23,156	0.115	0.319	0	1	UF
Agricultura	145,640	14,278	0.098	0.297	0	1	9907
Energia e mineração	145,640	4,286	0.029	0.169	0	1	9907
Indústria	145,640	11,991	0.082	0.275	0	1	9907
Indústria de alta tecnol.	145,640	5,390	0.037	0.189	0	1	9907
Serv. de utilidade pública	145,640	6,860	0.047	0.212	0	1	9907
Construção	145,640	12,191	0.084	0.277	0	1	9907
Comércio	145,640	33,957	0.233	0.423	0	1	9907
Transporte	145,640	6,946	0.048	0.213	0	1	9907
Serviços de alta tecnol.	145,640	24,848	0.171	0.376	0	1	9907
Outros serviços	145,640	15,887	0.109	0.312	0	1	9907
Setor público	145,640	9,006	0.062	0.241	0	1	9907
Sindicalizado	157,535	28,218	0.179	0.383	0	1	9087
Outra ocupação	145,640	137,188	0.942	0.234	0	1	9906
Chefe	145,640	2,065	0.014	0.118	0	1	9906
Gerente	145,640	6,387	0.044	0.205	0	1	9906
Área metropolitana	201,398	79,607	0.395	0.489	0	1	4107
Empregado	201,398	145,640	0.723	0.447	0	1	4718
Casado	182,181	150,261	0.825	0.380	0	1	4723
Mãe com criança	182,181	10,372	0.057	0.232	0	1	4723

*Nota:* SKC é a abreviatura para concentração setorial de conhecimento, o logaritmo da fração de pessoas em cada setor que possuam pelo menos um mestrado. KSI é a abreviatura para intensidade de difusão do conhecimento, o logaritmo da fração de pessoas em cada estado que estão empregadas no setor de serviços às empresas. Pessoa de refer. é a pessoa de referência da família de acordo com o respondente do questionário. As variáveis de setor são definidas na Tabela 5.2. Crianças são definidas como indivíduos com menos de 14 anos de idade.

*Fonte:* PNAD 2008.

adequado. O volume de 2008 contém um total de 391.816 observações que representam uma população de 190 milhões. Desse total, acabou-se usando apenas 201.398 observações de indivíduos com idade entre 20 e 60 anos, que representam 97,4 milhões de pessoas.

Do total de 201.398 observações, 145.640 (72,3%) pessoas estavam empregadas na semana de referência da pesquisa (21-27 set. 2008). A Tabela 5.1 mostra estatísticas so-

bre as variáveis utilizadas nas regressões mostradas adiante. O salário mensal individual tem uma média não ponderada de R\$ 1.015 e foi truncado em R\$ 30.000.<sup>3</sup> Os anos de escolaridade variam continuamente na faixa 0-15, cujo limite superior representa curso superior completo, e tem um valor adicional de 17, que representa as pessoas acima do ensino superior<sup>4</sup>. A escolaridade média é de 8,11 anos. As variáveis idade e experiência são mostrados em décadas. A experiência é a experiência no mercado de trabalho informada pelo indivíduo e, assim, *não* experiência potencial como usual.<sup>5</sup> A renda familiar possui os extremos R\$ 0 e R\$ 158.500. As famílias são compostas por 3,6 membros, em média, e têm entre zero e nove crianças com menos de 14 anos. A fração de pessoas em cada setor com escolaridade acima da graduação (SKC) varia de -9,545 (ou 0,007%) no setor 10 (serviços domésticos, reciclagem, comércio ambulante, limpeza de ruas e outras atividades) a -2,023 (ou 13,23%) no setor 9 (correios, telecomunicações, intermediação financeira, informática, pesquisa e desenvolvimento, educação, serviços às empresas, lazer e atividades associativas). A fração de pessoas em cada estado empregadas no setor de serviços às empresas (KSI) varia de -5,969 (ou 0,26%) no estado de Alagoas a -3,082 (ou 4,59%) no Distrito Federal. Para as variáveis categóricas, a Tabela 5.1 mostra o número não ponderado de observações com valor 1 na coluna “soma”, e a respectiva proporção percentual que representam no número total de observações não faltantes na coluna “média”. Assim, 50,9% da amostra é do sexo feminino, 0,3% é trabalhador militar, 53,8% é preto ou pardo, 4,7% é empregado no setor dos serviços, e 17,9% é sindicalizado. As variáveis de setor foram construídas a partir da Classificação Nacional de Atividades Econômicas (CNAE/IBGE), edição *Domiciliar*, e suas definições rigorosas são apresentadas na Tabela 5.2.

---

<sup>3</sup>Nas regressões na verdade foi utilizado o logaritmo do salário por hora trabalhada, computado a partir das perguntas 4718 (salário mensal) e 9058 (horas normalmente trabalhadas por semana) da pesquisa.

<sup>4</sup>O levantamento da PNAD não permite discriminar entre diferentes níveis de estudos de pós-graduação. A pergunta 4803 da pesquisa, de fato, assume valores da faixa 1-17, onde 1 significa menos de um ano de escolaridade, 2 significa um ano de escolaridade e assim por diante até o valor 16, que significa 15 anos ou mais, e 17, que significa não determinado. A variável educação foi, portanto, combinada com a questão 6003, que é sobre o nível de educação atual do indivíduo, e 6007, que é sobre o nível de educação mais alto que o indivíduo tenha alcançado.

<sup>5</sup>Estritamente falando, a **experiência** foi computada como a idade do indivíduo (pergunta 8005 da pesquisa) menos a idade em que ele informou ter começado a trabalhar (pergunta 9892), dividida por dez, para as pessoas que responderam a última pergunta; para os demais indivíduos, a experiência foi computada pela sua idade menos seus anos de escolaridade, dividido por dez, ou ausente, se o valor obtido foi menor que zero.

Tabela 5.2: Definição das variáveis de setor

Variável	Código CNAE	Descrição
Agricultura	11 12 13 14 20 50	Agropecuária e serviços relacionados, silvicultura e exploração florestal, pesca e aquicultura
Energia & mineração	100 110 230 234 130 140 260 270 280	Extração e processamento de materiais energéticos e minerais (excl. máquinas e equipamentos), e serviços relacionados
Indústria	150 160 170 180 190 200 360	Alimentos, têxteis, confecção, couros e calçados, madeira e móveis etc.
Indústria de alta tecnol.	210 220 240 250 290 300 310 320 330 340 350	Celulose, papel e edição, produtos químicos, plástico, máquinas e equipamentos, veículos etc.
Serv. de utilidade pública	400 410 850	Eletricidade, gás, água, saúde, serviços sociais etc.
Construção	459	Construção civil
Comércio	500 530 550 700 450 710 930	Comércio, reparação, alojamento, alimentação, imobiliárias, aluguel de móveis, serviços pessoais
Transporte	600 610 620 630	Transporte e atividades relacionadas
Serv. de alta tecnol.	640 650 660 670 720 730 740 800 910 920	Correio, telecomunicações, intermediação financeira, informática, pesquisa e desenvolvimento, educação, serviços às empresas, atividades associativas e recreativas
Outros serviços	370 531 900 950 990 998	Serviços domésticos, reciclagem, comércio ambulante, limpeza urbana e outras atividades
Setor público	750	Administração pública, defesa e seguridade social

*Nota:* Os códigos referem-se aos primeiros três dígitos da classificação da CNAE Domiciliar/IBGE. A codificação da CNAE Domiciliar pode ser obtida em <http://www.ibge.gov.br/concla/>.

*Fonte:* Elaboração própria.

## 6 Resultados

A Tabela 6.1 mostra os coeficientes estimados. A primeira coluna (MQO) mostra os coeficientes estimados por mínimos quadrados ordinários com erros-padrão corrigidos para heterocedasticidade pelo procedimento de White. Os coeficientes que figuram na segunda coluna (Heckman) foram estimados por máxima verossimilhança utilizando o procedimento desenvolvido por Heckman (1979), considerando também o desenho amostral complexo da PNAD, isto é, pesos de amostragem, amostragem por conglomerados e estratificação.

Embora as estatísticas apontem que tanto o desenho amostral complexo quanto o viés de não-seleção devam ser considerados, a Tabela 6.1 mostra que a maioria dos coeficientes não variam muito entre as especificações. Não obstante, os erros-padrão tendem a ser menores na especificação MQO. As maiores discrepâncias entre coeficientes foram observadas para as variáveis indicadoras (exceto *região metropolitana*) e para a variável KSI, cujo coeficiente é bem menor após os ajustes. A medida de ajuste  $r^2$  é de 0,463 para a especificação MQO.

Como tanto SKC quanto KSI são tomados em logaritmo, seus coeficientes nos dão a sua elasticidade do salário. Ambas as variáveis são estatisticamente significantes, e exercem um importante efeito sobre os salários. Olhando para a coluna Heckman, vê-se que um aumento de um por cento na fração de pessoas muito educadas (SKC) num determinado setor está associado com um aumento de (aproximadamente) 3,26% nos salários das pessoas empregadas nesse setor. O efeito da intensidade de difusão do conhecimento é ainda maior: um aumento de um por cento na fração estadual de pessoas envolvidas no setor de serviços às empresas está associado com um aumento de (aproximadamente) 4,45% nos salários das pessoas que trabalham nesse estado. Estes efeitos são, respectivamente, evidência para externalidades setoriais do capital humano e para o papel de difusão do conhecimento desempenhado pelo setor de serviços às empresas. Efeitos regionais sobre os salários já foram considerados por meio da inclusão de *dummies* de região na regressão.

A Tabela 6.1 também mostra os coeficientes da equação de seleção, que são significantes ao nível de 5%, exceto *mãe com criança* e as regiões Norte e Centro-Oeste (um teste de Wald revela que as *dummies* de região são conjuntamente significantes ao nível

Tabela 6.1: Regressão do logaritmo do salário

Variável	MQO	Heckman	Variável	MQO	Heckman
<b>Equação de salários</b>					
Escolaridade	0.08560 (0.00356)	0.07420 (0.00414)	Mulher	-0.19200 (0.00455)	-0.18900 (0.00526)
Escolarid. ao quadr.	-0.01190 (0.00051)	-0.01020 (0.00057)	Área rural	-0.12700 (0.00699)	-0.10600 (0.01070)
Escolarid. ao cubo	0.00078 (0.00002)	0.00066 (0.00002)	Pessoa de refer.	0.09830 (0.00406)	0.10200 (0.00443)
Experiência/10	0.21300 (0.00581)	0.21100 (0.00638)	Norte	0.01930 (0.00924)	-0.01270 (0.01960)
Exper./10 ao quadr.	-0.02930 (0.00123)	-0.02820 (0.00133)	Nordeste	-0.26800 (0.00574)	-0.27700 (0.01190)
SKC	0.02580 (0.00316)	0.03260 (0.00390)	Sul	0.02240 (0.00536)	-0.02900 (0.00936)
KSI	0.09370 (0.00705)	0.04450 (0.01400)	Centro-Oeste	0.07870 (0.00608)	0.04930 (0.01150)
Militar	0.28700 (0.03110)	0.29200 (0.03360)	Sindicalizado	0.11700 (0.00493)	0.10400 (0.00680)
Func. público	0.14000 (0.00776)	0.12600 (0.00945)	Chefe	0.48400 (0.02060)	0.43600 (0.02580)
Empregado informal	-0.13600 (0.00455)	-0.10600 (0.00571)	Gerente	0.28700 (0.01120)	0.27000 (0.01200)
Autônomo	-0.03640 (0.00609)	0.03740 (0.00858)	Área metropol.	0.09330 (0.00419)	0.09630 (0.00655)
Empregador	0.45400 (0.01300)	0.42500 (0.01580)	Crianças na família	-0.01850 (0.00192)	-0.02470 (0.00258)
Preto/pardo	-0.11400 (0.00394)	-0.13900 (0.00539)	Casado	0.09780 (0.00484)	0.06340 (0.00631)
Agricultura	-0.16500 (0.01060)	-0.13600 (0.01450)	Construção	0.03530 (0.00711)	0.01540 (0.00807)
Energia e mineração	0.17300 (0.01020)	0.15100 (0.01160)	Transporte	0.17800 (0.00933)	0.18100 (0.01020)
Indústria	-0.06620 (0.00703)	-0.05680 (0.00938)	Serv. de alta tecnol.	0.10300 (0.01090)	0.09080 (0.01330)
Indúst. de alta tecnol.	0.14200 (0.01010)	0.13700 (0.01260)	Outros serviços	0.00900 (0.01140)	0.04830 (0.01350)
Serv. de util. pública	0.15800 (0.01270)	0.10000 (0.01530)	Setor público	0.28300 (0.01210)	0.22200 (0.01470)
Constante	1.18000 (0.03720)	1.42000 (0.06380)			
N	145,640	179,899	r2	0.463	
F	4,186	1,122			
<b>Equação de seleção</b>					
Renda familiar		0.00014 (0.00001)	Casado	0.07190 (0.01110)	
Membros da família		-0.08050 (0.00378)	Mãe com criança	-0.00109 (0.01560)	
Crianças na família		0.10700 (0.00543)	Norte	0.01270 (0.01680)	
Escolaridade		0.00931 (0.00206)	Nordeste	0.03100 (0.01290)	
Idade/10		-0.01320 (0.00435)	Sul	0.13500 (0.01420)	
Preto/pardo		0.11500 (0.00864)	Centro-Oeste	0.01180 (0.01540)	
Constante		0.28000 (0.02640)			
$\rho$		-0.81700 (0.00698)	$\lambda$	-0.61576 (0.00823)	

Fonte: PNAD 2008.

de 5%). Os sinais dos coeficientes para membros da família e crianças na família são interessantes. Eles dizem que o salário de reserva tende a ser maior para membros de

famílias mais numerosas e a ser menor para indivíduos de famílias com muitas crianças. O salário de reserva também parece ser menor para pessoas casadas em relação aos outros. O Nordeste e o Sul parecem ter pessoas com um salário de reserva menor do que aquelas da região de referência Sudeste.

Tabela 6.2: Retornos de escala, %

MQO			Heckman		
Escolaridade	$\frac{\partial \ln(w)}{\partial S}$	$\frac{\partial \ln(w)^2}{\partial^2 S}$	Escolaridade	$\frac{\partial \ln(w)}{\partial S}$	$\frac{\partial \ln(w)^2}{\partial^2 S}$
4	2.8	-0.50	4	2.4	-0.46
8	4.6	1.38	8	3.8	1.13
11	10.8	2.79	11	9.0	2.32
15	25.7	4.66	15	21.4	3.91
17	36.0	5.60	17	30.0	4.71
5.1	2.5	0.00	5.1	2.2	0.00

Fonte: PNAD 2008.

Os coeficientes das variáveis de escolaridade não podem ser interpretados diretamente. A Tabela 6.2 mostra os retornos à educação mincerianos para cinco níveis de escolaridade para cada especificação, bem como os anos de escolaridade que marcam o início de retornos crescentes à escolaridade. Os cálculos são idênticos aos de Trostel (2004). Em primeiro lugar, os retornos à educação são consideravelmente mais baixos na especificação Heckman. Isto revela um considerável viés de não-seleção, já que a regressão que considera apenas o desenho amostral (não mostrada) não apresenta grande diferença em relação à especificação MQO. Em segundo lugar, os retornos à escolaridade (coluna  $\partial \ln(w)/\partial S$ ) são crescentes, passando de 2,4% com quatro anos de escolaridade a 30,0% com 17 anos de escolaridade ou mais, na especificação Heckman. Em terceiro lugar, os retornos crescentes à educação são atingidos a partir de 5,1 anos de estudo. As colunas denominadas de  $\partial \ln(w)^2/\partial^2 S$  mostram derivadas de segunda ordem do logaritmo do salário em relação à anos escolarização.

Quando analisam-se os retornos à educação por grandes setores<sup>1</sup> há claramente diferenças entre eles. A Tabela 6.3 mostra que os retornos à escolaridade são muito maiores para a indústria de alta tecnologia (por exemplo, papel, produtos químicos, plásticos, máquinas e equipamentos, veículos), mas também relativamente grandes para outras

<sup>1</sup>Mais especificamente, para estimar os retornos setoriais à escolarização, variáveis binárias de setor foram interagidas com as variáveis de escolaridade. A variável SKC é descartada por causa de colinearidade perfeita.

Tabela 6.3: Retornos setoriais à escolarização, %

Setor	N	4	8	11	15	17	$\frac{\partial \ln(w)^2}{\partial^2 S} = 0$	Escolaridade média	SKC
Agricultura	14,278	4.8	4.9	6.3	10.2	13.0	5.9	4.0	0.05
Energia e mineração	4,286	3.0	5.6	10.7	21.8	29.2	3.9	8.5	0.65
Indústria	11,991	3.3	3.4	8.8	23.0	33.2	5.9	7.8	0.15
Indústria de alta tecnol.	5,390	2.5	4.1	10.1	24.3	34.0	5.1	9.9	0.83
Serv. de utilidade pública	6,860	0.6	5.0	12.5	28.0	38.1	3.2	11.4	2.98
Construção	12,191	1.6	3.9	10.3	25.1	35.1	4.7	6.2	0.19
Comércio	33,957	2.8	4.0	7.2	14.7	19.8	4.7	8.8	0.18
Transporte	6,946	2.2	2.8	6.8	16.7	23.7	5.5	8.0	0.10
Serv. de alta tecnol.	24,848	1.5	5.7	10.3	18.3	23.1	-1.8	11.6	3.45
Outros serviços	15,887	2.0	1.0	2.1	5.9	8.8	7.5	6.1	0.02
Setor público	9,006	2.2	6.4	12.0	22.7	29.4	1.4	11.1	2.11
Todos os setores	145,640	2.4	3.8	9.0	21.4	30.0	5.1	8.1	0.99

Fonte: PNAD 2008.

*Nota:*  $N$  é o número não-ponderado de observações. A coluna  $\partial \ln(w)^2 / \partial^2 S = 0$  mostra anos de escolaridade mínimos requeridos para obter retornos crescentes de escala à escolaridade.  $SKC$  aqui é a fração percentual de pessoas com pelo menos um mestrado, e não o logaritmo dessa fração como no restante do texto.

indústrias (por exemplo, alimentos, têxtil, vestuário, madeira, mobiliário), serviços de utilidade pública, construção, serviços de alta tecnologia (por exemplo, correios, telecomunicações, intermediação financeira, informática, pesquisa e desenvolvimento, educação, serviços prestados às empresas) e setor público. Observam-se retornos mais baixos para os setores energia, agricultura e mineração, comércio, transportes e outros serviços (por exemplo, serviços domésticos, reciclagem, comércio ambulante, limpeza de ruas). Dito isso, notam-se retornos crescentes de escolaridade para todos os setores relatados. A maioria dos setores demonstra retornos crescentes a partir de cinco a seis anos de escolaridade, mas alguns deles mesmo antes. A análise por grandes setores não permite estabelecer nenhuma relação entre o estoque de capital humano por setor, medido pela escolaridade média e pela fração de pessoas com título de mestre ou superior a este (últimas duas colunas da tabela), e os retornos à escolaridade. A análise por setores mais desagregados não foi possível devido à perda de graus de liberdade. Os coeficientes das variáveis  $KSI$  e  $SKC$  continuam significantes a qualquer nível habitual de significância, e suas estimativas pontuais são 4,49% e 3,44%, respectivamente – não tão diferentes das estimativas anteriores.

## 7 Considerações finais

Efeitos de espraiamento do conhecimento, bem como externalidades tecnológicas e pecuniárias do capital humano certamente são importantes no processo de crescimento a longo prazo da economia. Os modelos de crescimento endógeno oferecem um suporte teórico para pensar sobre a relação de senso comum entre eles. Neste trabalho foram usadas ideias originadas a partir desses modelos para derivar uma equação de rendimento empiricamente estimável e para testar se há evidências de externalidades intra-setoriais do capital humano (tecnológicas ou pecuniárias) na economia brasileira, na tentativa de encontrar uma medida do efeito do “setor de pesquisa” de Romer (1990) e sobre os salários e, portanto, sobre a produtividade do trabalho.

Além disso, cabe ressaltar que se o conhecimento é ao mesmo tempo não-rival e um insumo de produção que melhora a produtividade dos outros fatores de produção (trabalho e capital, por exemplo), então empresas que lucram com a venda de informações produtivas e, portanto, conhecimento, deveriam obter retornos especialmente elevados. Pode-se identificar o setor de serviços às empresas como tendo exatamente essa característica. Supondo que os retornos mais elevados são pelo menos parcialmente transferidos para os empregados, foi testado o efeito deste setor sobre os salários individuais, em nível estadual.

Ao mesmo tempo, considerando que o setor de pesquisa seja composto por trabalhadores altamente qualificados, foi abordada uma questão importante apontada por Lucas (1988) quando ele argumenta em favor de importantes “efeitos externos” da acumulação de capital humano, embora ele estivesse preocupado principalmente com efeitos de interação pessoal limitados espacialmente e com efeitos de “aprender-fazendo” (*learning-by-doing*).

Como resultado, mostram-se evidências de que os efeitos de espraiamento possuem um traço setorial importante na economia brasileira. Usando equações de salário de tipo minceriano mostra-se o salário individual é influenciado positiva e significativamente pela concentração setorial de conhecimento (aproximado pela fração de pessoas do setor onde o indivíduo trabalha que possuam pelo menos o grau de mestre). O resultado se mantém após controlar para ocupação, etnia, gênero, área metropolitana, efeitos fixos regionais, entre outros. Também há evidência de que a intensidade de difusão do conhecimento

(aproximada pela fração de pessoas de um determinado estado empregadas no setor de serviços às empresas) exerça um efeito positivo e significativo sobre o salário dos indivíduos de determinado estado. Acredita-se que ambos os resultados suportem o modelo teórico de Romer (1990), bem como a existência de externalidades do capital humano na economia brasileira.

O valor mínimo encontrado após o qual surgem retornos crescentes à escolaridade (em geral cerca de 5,1 anos, e cerca de 4 a 6 anos para a maioria dos setores) é compatível com (embora um pouco maior) a evidência sobre o efeito limiar obtida por Dias e McDermott (2003), que estimam um valor limiar de três anos de escolaridade para o Brasil utilizando os dados da Pesquisa Mundial de Valor (*World Value Survey*).

Usando a fração setorial de pessoas altamente qualificadas (com grau de mestre ou superior a este), estimou-se uma externalidade da intensidade setorial de pesquisa. Um aumento de um por cento na referida fração leva a um aumento aproximado de 3,26% nos salários setoriais. Por sua vez, o efeito estadual de um aumento de um por cento na fração de pessoas empregadas no setor de serviços às empresas sobre os salários individuais é estimado em 4,45%. Estes resultados continuam válidos mesmo que se interajam as variáveis de escolaridade com as variáveis binárias de setor.

A terceira contribuição deste artigo é reforçar as evidências de retornos crescentes à escolaridade, ou seja, de retornos de escala na acumulação de um dos principais insumos da função de capital humano. Além disso, foi mostrado que esses retornos crescentes não são uma característica de qualquer setor específico, mas aparecem em maior ou menor grau em todos os grandes setores econômicos. No entanto, existem algumas diferenças importantes entre os setores em relação aos retornos à educação que ainda precisam ser explicadas. Retornos à escolaridade especialmente baixos são observados na agricultura e em outros serviços.

Há boas razões para se verificar melhor a existência e a magnitude de efeitos setoriais de transbordamento do conhecimento e das externalidades do capital humano. Não são apenas os desenvolvimentos teóricos que apontam que esses efeitos provavelmente existam e sejam importantes para compreender o processo de desenvolvimento econômico, mas esses efeitos também são ainda pouco explorados no Brasil.

## Referências

ACEMOGLU, D. A microfoundation for social increasing returns in human capital accumulation. *The Quarterly Journal of Economics*, v. 111, n. 3, p. 779–804, ago. 1996.

ACEMOGLU, D. *Introduction to Modern Economic Growth*. Princeton, NJ: Princeton University Press, 2008.

ACEMOGLU, D.; ANGRIST, J. How large are human-capital externalities? evidence from compulsory-schooling laws. In: BERNANKE, B. S.; ROGOFF, K. (Ed.). *NBER Macroeconomics Annual 2000, Vol. 15*. Cambridge: MIT Press, 2001, (NBER Chapters). p. 9–74.

ANTELIUS, J. *Sheepskin Effects in the Returns to Education: Evidence on Swedish Data*. Stockholm: Trade Union Institute for Economic Research, 2000. (FIEF Working Paper Series, 158/2000).

ARROW, K. J. The economic implications of learning by doing. *The Review of Economic Studies*, v. 29, n. 3, p. 155–173, jun. 1962.

BARRO, R. T.; SALA-I-MARTIN, X. *Economic Growth*. 2. ed. Cambridge, Massachusetts; London, England: MIT Press, 2004.

BECKER, G. S. Investment in human capital: A theoretical analysis. *The Journal of Political Economy*, v. 70, n. 5, part 2, p. 9–49, out. 1962.

BEHRMAN, J. R.; BIRDSALL, N. The quality of schooling: Quantity alone is misleading. *The American Economic Review*, v. 73, n. 5, p. 928–946, dez. 1983.

BELMAN, D.; HEYWOOD, J. S. Sheepskin effects by cohort: Implications of job matching in a signaling model. *Oxford Economic Papers*, v. 49, n. 4, p. 623–637, 1997.

CARD, D.; KRUEGER, A. B. Does school quality matter? returns to education and the characteristics of public schools in the united states. *The Journal of Political Economy*, v. 100, n. 1, p. 1–40, fev. 1992.

- CASS, D. Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies*, v. 32, n. 3, p. 233–240, jul. 1965.
- CHARLOT, S.; DURANTON, G. Communication externalities in cities. *Journal of Urban Economics*, v. 56, n. 3, p. 581–613, nov. 2004.
- CICCONE, A.; PERI, G. Identifying human-capital externalities: Theory with applications. *The Review of Economic Studies*, v. 73, n. 2, p. 381–412, abr. 2006.
- SOCIETY, B. E. (Ed.), XXV., 2003, Porto Seguro, BA. *Aggregate Threshold Effects and the Importance of Human Capital in Economic Development*. Porto Seguro: Brazilian Meeting of Econometrics, 2003.
- DUFLO, E. The medium run effects of educational expansion: Evidence from a large school construction program in indonesia. *Journal of Development Economics*, v. 74, n. 1, p. 163–197, jun. 2004.
- EASTERLY, W. et al. *Policy, Technology Adoption, and Growth*. [S.l.]: National Bureau of Economic Research, 1994. (NBER Working Papers, 4681).
- FLORES-LAGUNES, A.; LIGHT, A. *Identifying Sheepskin Effects in the Returns to Education*. Tucson, AZ: Department of Economics, University of Arizona, 2004. (Working Paper).
- FU, S. Smart café cities: Testing human capital externalities in the boston metropolitan area. *Journal of Urban Economics*, v. 61, n. 1, p. 86–111, jan. 2007.
- HANUSHEK, E. A. The economic value of education and cognitive skills. In: SYKES, G.; SCHNEIDER, B.; PLANK, D. N. (Ed.). *Handbook of Education Policy Research*. New York: Routledge, 2009. cap. 3, p. 39–56.
- HANUSHEK, E. A. Education production functions: Developed country evidence. In: BAKER, E.; MCGAW, B.; PETERSON, P. (Ed.). *International Encyclopedia of Education*. 3. ed. Oxford: Elsevier, 2010. p. 407–411.
- HANUSHEK, E. A.; KIMKO, D. D. Schooling, labor-force quality, and the growth of nations. *The American Economic Review*, v. 90, n. 5, p. 1184–1208, 2000.

HECKMAN, J.; LAYNE-FANAR, A.; TODD, P. Human capital pricing equations with an application to estimating the effect of schooling quality on earnings. *The Review of Economics and Statistics*, v. 78, n. 4, p. 562–610, nov. 1996.

HECKMAN, J. J. Sample selection bias as a specification error. *Econometrica*, v. 47, n. 1, p. 153–161, 1979.

HEUERMANN, D. F. *Human Capital Externalities in Western Germany*. Trier: Institute of Labour Law and Industrial Relations in the European Community (IAAEG), 2009. (Discussion Papers, 03/2009).

HEYNEMAN, S. P.; LOXLEY, W. A. The effect of primary-school quality on academic achievement across twenty-nine high-and low-income countries. *American Journal of Sociology*, v. 88, n. 6, p. 1162–1194, mai. 1983.

HOFFMANN, R.; SIMÃO, R. C. S. Determinantes do rendimento das pessoas ocupadas em minas gerais em 2000: O limiar no efeito da escolaridade e as diferenças entre mesorregiões. *Nova Economia*, v. 15, n. 2, p. 35–62, mai./ago. 2005.

JACOBS, J. *The economy of cities*. New York: Random House, 1969.

JAEGER, D. A.; PAGE, M. E. Degrees matter: New evidence on sheepskin effects in the returns to education. *The Review of Economics and Statistics*, v. 78, n. 4, p. 733–740, nov. 1996.

JAFFE, A. B.; TRAJTENBERG, M.; HENDERSON, R. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, v. 108, n. 3, p. 577–598, ago. 1993.

JONES, C. I. R&D-Based Models of Economic Growth. *Journal of Political Economy*, v. 103, n. 4, p. 759–784, ago. 1995.

JONES, C. I. *Introdução à Teoria do Crescimento Econômico*. Rio de Janeiro: Campus, 2000.

KOOPMANS, T. C. On the concept of optimal economic growth. *Pontificiae Academiae Scientiarum Scripta Varia*, North Holland, Amsterdam, v. 28, n. 1, p. 226–300, 1965.

- LAM, D.; SCHOENI, R. F. Effects of family background on earnings and returns to schooling: Evidence from Brazil. *Journal of Political Economy*, v. 101, n. 4, p. 710–740, ago. 1993.
- LUCAS, R. E. On the mechanics of economic development. *Journal of Monetary Economics*, v. 22, n. 1, p. 3–42, jul. 1988.
- MANKIW, G. The growth of nations. *Brookings Papers on Economic Activity*, v. 26, n. 1995-1, p. 275–326, 1995.
- MANKIW, N. G.; ROMER, D.; WEIL, D. N. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, v. 107, n. 2, p. 407–437, mai. 1992.
- MARSHALL, A. *Princípios de Economia*. 8. ed. London: Macmillan and Co., 1890.
- MCDERMOTT, J. Mercantilism and modern growth. *Journal of Economic Growth*, v. 4, n. 1, p. 55–80, mar. 1999.
- MCDERMOTT, J. Development dynamics: Economic integration and the demographic transition. *Journal of Economic Growth*, v. 7, n. 4, p. 371–409, dez. 2002.
- MINCER, J. Investment in human capital and personal income distribution. *The Journal of Political Economy*, v. 66, n. 4, p. 281–302, ago. 1958.
- MINCER, J. On-the-job training: Costs, returns, and some implications. *The Journal of Political Economy*, v. 70, n. 5, part 2, p. 50–79, out. 1962.
- MINCER, J. *Schooling, Experience, and Earnings*. New York: NBER; Columbia University Press, 1974.
- MONTEIRO, W. F.; DIAS, J.; DIAS, M. H. A. Taxa de retorno da escolaridade nos estados brasileiros: Crescente ou decrescente. In: *XXXVII Encontro Nacional de Economia*. Foz do Iguaçu: Rio de Janeiro: ANPEC, 2009. (Anais eletrônicos... , v. 37), p. 20 p.
- MORETTI, E. Estimation the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, v. 121, n. 1-2, p. 175–212, jul./ago. 2004.

- RAMOS, L. A.; VIEIRA, M. L. A relação entre educação e salários no Brasil. In: *A Economia Brasileira em Perspectiva: 1996*. Rio de Janeiro: IPEA, 1996. v. 2, cap. 21, p. 493–510.
- RAUCH, J. E. Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics*, v. 34, n. 3, p. 380–400, nov. 1993.
- ROMER, P. M. Increasing returns and long-run growth. *Journal of Political Economy*, v. 94, n. 5, p. 1002–1037, out. 1986.
- ROMER, P. M. Endogenous technological change. *The Journal of Political Economy*, v. 98, n. 5, p. S71–S102, out. 1990.
- ROSEN, S. A theory of life earnings. *The Journal of Political Economy*, v. 84, n. 4, p. S45–S67, ago. 1976.
- ROSENTHAL, S. S.; STRANGE, W. C. The attenuation of human capital spillovers. *Journal of Urban Economics*, v. 64, n. 2, p. 373–389, set. 2008.
- SALVATO, M. A.; SILVA, D. G. O impacto da educação nos rendimentos do trabalhador: Uma análise para região metropolitana de Belo Horizonte. In: *Anais do XIII Seminário sobre a Economia Mineira*. [S.l.]: Cedeplar, Universidade Federal de Minas Gerais, 2008. p. 23 p.
- SCHULTZ, T. W. Investment in human capital. *The American Economic Review*, v. 51, n. 1, p. 1–17, mar. 1961.
- SHELL, K. Toward a theory of inventive activity and capital accumulation. *The American Economic Review*, v. 56, n. 1/2, p. 62–68, mar. 1966.
- SOLOW, R. M. A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, v. 70, n. 1, p. 65–94, fev. 1956.
- SWAN, T. W. Economic growth and capital accumulation. *Economic Record*, v. 32, n. 2, p. 334–361, nov. 1956.
- TROSTEL, P. A. Returns to scale in producing human capital from schooling. *Oxford Economic Papers*, v. 56, n. 3, p. 461–484, jul. 2004.
- UZAWA, H. Optimum technical change in an aggregative model of economic growth. *International Economic Review*, v. 6, n. 1, p. 18–31, jan. 1965.

---

WHEATON, W. C.; LEWIS, M. J. Urban wages and labor market agglomeration.  
*Journal of Urban Economics*, v. 51, n. 3, p. 542–562, mai. 2002.

WHEELER, C. H. Do localization economies derive from human capital externalities?  
*Annals of Regional Science*, v. 41, n. 1, p. 31–50, mar. 2007.

## A Modelo de crescimento de Solow (1956)

Suponha que houvesse uma família representativa com dadas preferências, assim como uma empresa representativa com uma dada função de produção que vende um único bem às famílias. Nenhuma outra mercadoria, sem indivíduos ou organizações heterogêneas. Já que é representativa, a função de produção pode facilmente ser agregada para representar a economia inteira, principalmente se – como supomos – houver retornos constantes de escala. Especificamente, a função de produção agregada,

$$Y(t) = F[K(T), L(T), A(T)], \quad (\text{A.1})$$

depende do estoque de capital no tempo  $t$ ,  $K(t)$ , do emprego total  $L(t)$ , e da tecnologia  $A(t)$ . A “tecnologia” é não-rival e não-excluível, e incorpora não apenas o progresso técnico, mas também fatores organizacionais e bens públicos básicos como a garantia do direito à propriedade e a segurança nacional. A função de produção toma argumentos não negativos que correspondem a níveis não negativos de produto. Ela também é diferenciável continuamente por duas vezes em  $K$  e  $L$  (o argumento de tempo é omitido por conveniência), e satisfaz

$$F_K(K, L, A) \equiv \frac{\partial F(K, L, A)}{\partial K} > 0, \quad F_L(K, L, A) \equiv \frac{\partial F(K, L, A)}{\partial L} > 0, \quad (\text{A.2})$$

$$F_{KK}(K, L, A) \equiv \frac{\partial^2 F(K, L, A)}{\partial K^2} < 0, \quad F_{LL}(K, L, A) \equiv \frac{\partial^2 F(K, L, A)}{\partial L^2} < 0. \quad (\text{A.3})$$

As derivadas de segunda ordem negativas asseguram que ambos os insumos exibam retornos decrescentes. Como já observado, a função exibe retornos constantes de escala, isto é, ela é homogênea de grau um e, portanto, côncava. A homogeneidade linear implica (pelo teorema do Euler) que a soma dos produtos marginais do capital e o do trabalho exaurem o produto.

Supõe-se que todos os mercados de fatores sejam competitivos e estejam em equilíbrio geral. As famílias ofertam o seu trabalho inelasticamente para qualquer salário positivo [ $w(t) > 0$ ]. A condição de equilíbrio (ou compensação) do mercado de trabalho é

$$L(t) = \bar{L}(t), \quad (\text{A.4})$$

onde  $L(t)$  indica a demanda por trabalho e assim o nível de emprego. Além do trabalho, as famílias provêem também capital para as empresas, que pagam por ele um preço de aluguel de  $R(t)$ . A condição de compensação de mercado é obtida igualando a oferta de capital pelas famílias e a sua demanda pelas empresas:  $K^s(t) = K^d(t)$ . Supõe-se que tanto as famílias quanto as empresas estejam otimizando, respectivamente, a sua utilidade e o seu lucro. A dotação inicial de capital é dada simplesmente por  $K(0)$ . Ademais, as famílias poupam uma fração constante  $s$  da sua renda que se transforma em capital. O preço do bem final,  $P(t)$ , é normalizado para 1, e há uma taxa de juro intertemporal de  $r(t)$ . Estas últimas condições são especificadas para permitir um equilíbrio geral relativamente simples, como explica Acemoglu (2008, p. 37-38). O capital se desvaloriza (exponencialmente) à taxa  $\delta$ , a qual dá a relação entre a taxa de juro e a taxa de aluguel do capital:  $r(t) = R(t) - \delta$ . É importante ter em mente que a mesma unidade de produto (ou aluguel) pode ou ser consumida ou ser poupada e, assim, transformada em capital.

As empresas enfrentam o seguinte problema de otimização:

$$L(t) \geq 0, \bar{K}(t) \geq 0 \max F[K(t), L(t), A(t)] - w(t)L(t) - R(t)K(t), \quad (\text{A.5})$$

com condições necessárias de primeira-ordem

$$w(t) = F_L[K(t), L(t), A(t)] \quad (\text{A.6})$$

e

$$R(t) = F_K[K(t), L(t), A(t)]. \quad (\text{A.7})$$

O teorema de Euler implica ademais que as empresas obtêm lucro e que

$$Y(t) = w(t)L(t) + R(t)K(t). \quad (\text{A.8})$$

As seguintes condições adicionais de contorno (*condições de Inada*) asseguram a existência de equilíbrios interiores e implicam que as primeiras unidades de capital e trabalho sejam extremamente produtivas e que os seus produtos marginais tendam a zero a medida que suas quantidades consumidas aumentem:

$$\begin{aligned} \lim_{K \rightarrow 0} F_K(K, L, A) &= \infty \text{ e} \\ \lim_{K \rightarrow \infty} F_K(K, L, A) &= 0 \text{ para todo } L > 0 \text{ e todo } A; \\ \lim_{L \rightarrow 0} F_L(K, L, A) &= \infty \text{ e} \\ \lim_{L \rightarrow \infty} F_L(K, L, A) &= 0 \text{ para todo } K > 0 \text{ e todo } A. \end{aligned} \quad (\text{A.9})$$

Em outras palavras, os insumos da função de produção são *essenciais* mas ineficazes em gerar crescimento infinito.

Outros pressupostos comuns vêm da contabilidade nacional de uma economia fechada e sem governo:

$$Y(t) = C(t) + I(t) \quad (\text{A.10})$$

$$S(t) = I(t) = Y(t) - C(t) \quad (\text{A.11})$$

$$S(t) = sY(t) \quad (\text{A.12})$$

$$C(t) = (1 - s)Y(t), \quad (\text{A.13})$$

onde  $Y$  é produto,  $C$  consumo,  $I$  investimento,  $S$  poupança e  $s$  é uma fração constante da renda que é poupada pelas famílias.

Como o capital desvaloriza à taxa  $\delta$ , o estoque de capital no período  $t + 1$  é dado estoque de capital não-depreciado do período  $t$  mais os investimentos adicionais:

$$K(t + 1) = (1 - \delta)K(t) + I(t). \quad (\text{A.14})$$

Usando as relações da contabilidade nacional (A.13) chega-se à equação

$$K(t + 1) = (1 - \delta)K(t) + sY(t) \quad (\text{A.15})$$

$$K(t + 1) = sF[K(t), L(t), A(t)] + (1 - \delta)K(t), \quad (\text{A.16})$$

que é conhecida como a *lei fundamental do movimento* do modelo de Solow-Swan.

A proporção capital-produto e o produto (ou a renda) per capita são definidos, respectivamente, como

$$k(t) = \frac{K(t)}{L} \quad (\text{A.17})$$

Com retornos constantes de escala, e supondo uma população positiva e constante, bem como a ausência de progresso tecnológico, o segundo pode ser expresso por

$$y(t) = F\left(\frac{K(t)}{L}, 1, A\right) \equiv f[k(t)], \quad (\text{A.18})$$

ou seja, o produto per capita simplesmente como uma função da proporção capital-trabalho.

O produto marginal e o preço de aluguel do capital são dados pela derivada de  $F$  com relação ao seu primeiro argumento, que é  $f'(k)$ . O produto marginal do trabalho e a taxa de salário são então obtidos pelo teorema de Euler:

$$R(t) = f'[k(t)] \quad (\text{A.19})$$

$$w(t) = f[k(t)] - k(t)f'[k(t)]. \quad (\text{A.20})$$

Dividindo ambos os lados de (A.16) por  $L$ , obtemos a *equação diferencial de equilíbrio* do modelo de Solow, a qual descreve o comportamento de equilíbrio da variável-chave do modelo, a proporção capital-trabalho:

$$k(t+1) = sf[k(t)] + (1 - \delta)k(t). \quad (\text{A.21})$$

No equilíbrio de estado estável, todas as variáveis crescem a uma taxa constante. Dado que não há crescimento demográfico no presente modelo ( $\dot{L} = 0$ ), isto implica que o nível da razão de capital também deve permanecer constante. Igualando  $k(t+1) = k(t)$ , (A.21) se torna

$$\begin{aligned} k^*(t) &= sf[k^*(t)] + (1 - \delta)k^*(t) \\ k^*(t) - k^*(t) + \delta k^*(t) &= sf[k^*(t)] \\ \frac{f[k^*(t)]}{k^*(t)} &= \frac{\delta}{s} \end{aligned} \quad (\text{A.22})$$

onde  $k^*$  indica o valor em estado estável de  $k$ .

Se for permitido o crescimento demográfico, a condição de estado estável modifica-se ligeiramente para

$$f[k^*(t)]/k^*(t) = n + \delta/s, \quad (\text{A.23})$$

e a equação diferencial de equilíbrio (A.21) para

$$k(t+1) = sf[k(t)] - (n + \delta)k(t). \quad (\text{A.24})$$

No modelo do tipo Solow com progresso tecnológico é recomendado (para fins de simplicidade) expressar a função de produção em termos de variáveis normalizadas. Trabalho “efetivo” é definido como unidades de eficiência de trabalho, isto é,  $A(t)L(t)$ . Como  $F$  exhibe retornos constantes de escala nos seus dois argumentos, podemos definir

$$k(t) \equiv \frac{K(t)}{A(t)L(t)} \quad (\text{A.25})$$

como a *relação capital-trabalho efetiva*. Isto nos permite escrever a renda per capita  $y(t) \equiv Y(t)/L(t)$  como uma função da tecnologia  $e$  do capital por trabalho efetivo:

$$y(t) = A(t)f[k(t)]. \quad (\text{A.26})$$

Com progresso tecnológico, a lei de movimento de Solow se torna (em tempo contínuo)

$$\frac{\dot{k}(t)}{k(t)} = \frac{sf[k(t)]}{k(t)} - (\delta + g + n), \quad (\text{A.27})$$

onde  $g$  é a taxa positiva de crescimento tecnológico  $\dot{A}/A$ . Por sua vez, a condição de crescimento balanceado (de estado estável) passa a ser

$$\frac{f(k^*)}{k^*} = \frac{\delta + g + n}{s}. \quad (\text{A.28})$$

A dinâmica da trajetória de equilíbrio e da estática comparativa é muito similar àquela do modelo de Solow sem progresso tecnológico.

Introduzindo capital humano no modelo de Solow com mudança tecnológica tem-se uma função de produção da forma

$$Y = F(K, H, AL), \quad (\text{A.29})$$

onde  $H$  indica capital humano e atende às mesmas restrições de primeira e segunda ordem que o trabalho e o capital físico. Também as condições de Inada são adaptadas para incluir o capital humano, tratando-o exatamente como se fosse capital físico. Em muitos aspectos o capital humano pode ser tratado simplesmente como “capital”, supondo, por exemplo, que as pessoas poupem uma fração constante  $s_h$  de sua renda a qual investem na acumulação de capital humano, ou que o capital humano deprecia a uma taxa constante  $\delta_h$ .

Definindo a relação capital humano-trabalho efetiva como

$$h(t) = \frac{H(t)}{A(t)L(t)} \quad (\text{A.30})$$

pode-se obter a lei de movimento como anteriormente, com a diferença de que agora se tem duas:

$$\dot{k}(t) = s_k f[k(t), h(t)] - (\delta_k + g + n)k(t), \quad (\text{A.31})$$

$$\dot{h}(t) = s_h f[k(t), h(t)] - (\delta_h + g + n)h(t). \quad (\text{A.32})$$

O equilíbrio em estado estável, por sua vez, é único [veja Acemoglu (2008, cap. 3.3)] e dado por

$$s_k f(k^*, h^*) - (\delta_k + g + n)k^* = 0, \quad (\text{A.33})$$

e

$$s_h f(k^*, h^*) - (\delta_h + g + n)h^* = 0. \quad (\text{A.34})$$

Assumindo uma especificação Cobb-Douglas para a função de produção é possível escrever a trajetória de crescimento em estado estável da renda per capita do país  $j$  como

$$y_j^*(t) = A_j(t) \left( \frac{s_{k,j}}{n_j + g_j + \delta_k} \right)^{\frac{\alpha}{1-\alpha-\beta}} \left( \frac{s_{h,j}}{n_j + g_j + \delta_h} \right)^{\frac{\beta}{1-\alpha-\beta}}, \quad (\text{A.35})$$

onde o nível tecnológico é dado por  $A_j(t) = \bar{A}_j \exp(gt)$ . Isto é, os países diferem quanto ao seu nível tecnológico inicial, mas não quanto à sua taxa de crescimento. Este pressuposto pode assumir várias formas alternativas, sendo feito desta maneira por Mankiw, Romer e Weil (1992).

## B Modelo de crescimento de Romer (1990)

A economia no modelo de Romer consiste em três setores: bens finais, bens intermediários (de capital), pesquisa. O funcionamento básico da economia se dá como segue: O setor de pesquisa gera ideias que podem ser transformadas em bens de capital, e vende o direito de produzi-los a uma companhia no setor de bens intermediários; o último se torna um monopolista na produção desses bens de capital e os vende ao setor de bens finais que, por sua vez, gera o produto. Ademais, há quatro insumos na economia: capital físico (unidades de bens de consumo); força de trabalho (número de pessoas); capital humano (anos de instrução e treinamento formais); conhecimento ou tecnologia (número de “projetos” de novos bens de capital). O setor de pesquisa usa capital humano e o estoque de conhecimento existente para produzir projetos de novos bens de capital. O setor de bens intermediários usa os projetos criados pelo setor de pesquisa e o produto sacrificado para produzir um grande número de bens de capital, disponíveis para produção de bens finais. O setor de bens finais combina força de trabalho, capital humano e os bens de capital à sua disposição para produzir o produto final. O produto pode ser consumido ou poupado como novo capital.

Pressupostos simplificadores incluem: a população e a oferta de trabalho são constantes; o estoque total de capital humano na população é constante, assim como a sua parte destinada ao mercado.

O produto final  $Y$  é expresso como uma função da força de trabalho,  $L$ , do capital humano destinado ao produto final,  $H_Y$ , e do capital físico, expresso por bens de capital comumente disponíveis,  $\int_{i=1}^{\infty} x_i$ . Supondo que a função de produção tenha a forma de uma função Cobb-Douglas, obtém-se

$$Y(H_Y, L, x) = H_Y^\alpha L^\beta \int_{i=1}^{\infty} x_i^{1-\alpha-\beta} di. \quad (\text{B.1})$$

O que distingue esta função de outras funções de produção habituais é o fato de que os insumos de capital não são necessariamente substitutos e, portanto, não necessariamente possuem as mesmas produtividades marginais.

Enquanto o setor de bens finais pode ser descrito em termos de uma firma representativa e de um tomador de preço agregado, o setor de bens de capital não pode, porque há uma empresa específica para cada tipo de bem de capital produzido a partir de um projeto específico proveniente do setor de pesquisa – o que teoricamente pode ocorrer dentro da mesma empresa. Uma vez de posse do projeto do novo bem de capital, a empresa pode desviar (ou usar) o produto final para produzir esse bem.

Cada bem de capital é produzido por uma firma individual, tendo uma patente permanente que lhe permite obter uma renda  $p(i)$  das empresas de bens finais. Este arranjo dá poder de monopólio à empresa do setor de bens de capital e implica uma curva da demanda negativamente inclinada. Como não há depreciação, o valor de uma unidade do bem de capital  $i$  é o valor presente descontado do fluxo de renda infinito gerado por ele. Supondo que seja necessário abrir mão de  $\eta$  unidades de consumo para criar uma unidade de algum tipo de bem de capital, os bens de capital usados de fato na produção são dados por:

$$K = \eta \sum_{i=1}^{\infty} x_i = \eta \sum_{i=1}^A x_i. \quad (\text{B.2})$$

Logo,  $H$  e  $L$  são fixos, e  $K$  cresce proporcionalmente ao consumo sacrificado. Falta definir o processo de acumulação de novos projetos, isto é,  $\dot{A}$ . Para todos os pesquisadores,  $\sum^j(H)$ ,

a acumulação do estoque de projetos (isto é, do conhecimento tecnológico) é dado por

$$\dot{A} = \delta H_A A, \quad (\text{B.3})$$

Onde  $H_A$  é a proporção do capital humano empregado no setor de pesquisa. Dois importantes pressupostos desta função são: (i) a devoção de mais capital humano à pesquisa aumenta a taxa de invenção de novos projetos; (ii) quanto mais elevado for o estoque de projetos e conhecimento, maior é a produtividade de um engenheiro que trabalha na departamento de pesquisa. A linearidade em  $H_A$  não é essencial mas facilita a análise. A linearidade em  $A$  é o pressuposto que torna possível o crescimento ilimitado. Se  $A$  entrasse em (B.3) através de uma função côncava, o capital humano usado no setor de pesquisa se deslocaria para o setor industrial à medida que aumentasse  $A$ , o que conduziria sucessivamente a uma redução na taxa de crescimento, algo que não é suportado pela evidência empírica.

A característica crucial dessa especificação é que o conhecimento afeta a produção de duas formas. Um novo projeto permite a produção de um novo bem de capital, o qual pode ser usado para gerar produto. Um novo projeto também eleva o estoque de conhecimento e, conseqüentemente, a produtividade do capital humano no setor de pesquisa. O proprietário de um projeto tem direitos de propriedade sobre o seu uso na produção de bens de capital mas não sobre o seu uso em pesquisa. Neste sentido, o conhecimento é parcialmente excluível.

Por causa da simetria do modelo, todos os bens de capital disponíveis são ofertados na mesma quantidade, denotada por  $\bar{x}$ . Agora pode-se reescrever (B.2) como  $K = \eta A \bar{x}$ . Resolvendo para  $\bar{x}$  tem-se  $\bar{x} = K/\eta A$ . Substituindo este resultado em (B.1), obtém-se:

$$\begin{aligned} Y(H_Y, L, x) &= H_Y^\alpha L^\beta \int_0^\infty x_i^{1-\alpha-\beta} di \\ &= H_Y^\alpha L^\beta A \bar{x}^{1-\alpha-\beta} \\ &= H_Y^\alpha L^\beta A \left( \frac{K}{\eta A} \right)^{1-\alpha-\beta} \\ &= (H_Y A)^\alpha (L A)^\beta (K)^{1-\alpha-\beta} \eta^{\alpha+\beta-1}. \end{aligned} \quad (\text{B.4})$$

A não-convexidade evidente nesta equação é o resultado do equilíbrio de competição monopolista. Embora a origem dos retornos não-decrescentes seja o efeito de espraiamento do conhecimento,  $A$  afeta somente indiretamente o setor de bens finais por causa dos seus efeitos sobre a disponibilidade de novos bens de capital. No setor de pesquisa,  $A$  entra diretamente.

Como neste modelo uma fração constante da população é empregada na geração de ideias, a taxa de crescimento da renda per capita ao longo de uma trajetória de crescimento equilibrado é igual à taxa de crescimento da proporção capital-trabalho e igual à taxa de crescimento do estoque de ideias, como no modelo de Solow. A diferença em relação ao último é que no modelo de Romer a taxa de crescimento do estoque de conhecimento é endógena e dada por

$$g_A = \delta H_A. \quad (\text{B.5})$$

Romer supõe que a produtividade da pesquisa seja proporcional ao estoque existente de ideias:  $\bar{\delta} = \delta A$ . Com este pressuposto, a produtividade dos pesquisadores cresce no tempo ainda que o número de pesquisadores seja constante.

# C Rotina do Stata

```
// D:\DISSERTATION\DO\00-MASTER.do
////////////////////////////////////
* MASTER DO-FILE
* BASIC SETTINGS
capture log close
clear all
set mem 800m
set matsize 1000
set more off
global userdir "D:\DISSERTATION\DATA\"
global datadir "D:\DATA\PNAD\2008\"
* UNZIP & MERGE PES & DOM TO PNAD.DTA
do 01-MERGE.do
* EXTRACT DATA FROM PNAD.DTA
do 02-EXTRACT.do
* DESCRIBE RAW DATA (LOG1)
cd $userdir
do 03-LOG1-RAWDATA.do
* SAVE DATA
cd $userdir
save DADOS-BEFOREBASICVARS, replace
* CREATE BASIC VARIABLES
do 04-BASICVARS.do
* SAVE DATA
cd $userdir
save DADOS-BEFORESECTORVARS, replace
* CREATE SECTOR VARIABLES
do 05-SECTORVARS.do
* SAVE DATA
cd $userdir
save DADOS-BEFOREHCEVARS, replace
* CREATE VARIABLES TO MEASURE HUMAN CAPITAL EXTERNALITIES
do 06-HCEVARS.do
* RENAME SOME VARIABLES
do 07-RENAME.do
* ALLOCATE SINGLE PSU STRATA INTO LARGER STRATA
* Not necessary because treating strata with single sampling
* units as certainty units does not affect any result.
//do 08-SINGLEPSU.do
* DROP SOME OBSERVATIONS
do 09-DROP.do
* SAVE DATA
cd $userdir
save DADOS, replace
* DEFINE GLOBALS
global idvars "year state control series order wgt upa str psu"
global cvars "w educ age10 exp finc nfmem nchild knwconc knwsprd"
global ivars "fem i.position black rural reppers i.region i.mjsector union ///
i.occup metro works married mthchild"
global T "lwph educ educsq educcb exp expsq"
global X "ib(freq).position black fem rural reppers ib(freq).region union ///
ib(freq).occup ib(freq).mjsector metro nchild married"
global s "knwconc"
global r "knwsprd"
```

```

global H "finc nfmem nchild educ age10 i.region black married mthchild"
* DESCRIBE USED DATA (LOG2)
do 10-LOG2-USEDATA.do
* MAKE GRAPHS AND EXPORT TO PNG
do 11-GRAPHS.do
* MEASURE EFFECTS OF COMPLEX DESIGN AND NON-SELECTION
* ON COEFFICIENTS AND RATES OF RETURN TO SCHOOLING
* AND MEASURE SECTORAL EXTERNALITIES (LOG3)
do 12-LOG3-RETURNS.do
* MEASURE SECTORAL RETURNS TO SCHOOLING
* WITH POOLED REGRESSION (LOG4)
do 13-LOG4-POOLED.do
* MEASURE SECTORAL RETURNS TO SCHOOLING
* WITH SECTORAL REGRESSIONS (LOG5)
do 14-LOG5-SECTORS.do
* SAVE DATA
cd $userdir
save DADOS-FINAL, replace
// D:\DISSERTATION\DO\01-MERGE.do
////////////////////////////////////
* UNZIP & MERGE PES & DOM TO PNAD.DTA
version 10
cd $datadir
unzipfile DOM2008 , replace
infile using DCT-DOM2008.dct , clear
keep if V0104==1 // keep if interview realized
drop if mi(V0102,V0103) // drop if uniq. identif. var.s are miss.
isid V0102 V0103 // verify if var.s really uniq. identify obs.
sort V0102 V0103 , stable
save DOM2008 , replace
unzipfile PES2008 , replace
infile using DCT-PES2008.dct , clear
drop if mi(V0102,V0103,V0301) // drop if uniq. identif. var.s are miss.
isid V0102 V0103 V0301 // verify if var.s really uniq. identify obs.
sort V0102 V0103 V0301 , stable
save PES2008 , replace
merge V0102 V0103 using DOM2008 , uniquising
tab _merge // verify if _merge==3 for all obs.
drop if _merge!=3
drop _merge
save PNAD2008 , replace
/* PNAD2008 contains PES & DOM var.s and uniq. identif. obs. from
realized interviews */
erase DOM2008.dta
erase PES2008.dta
erase DOM2008.txt
erase PES2008.txt
// D:\DISSERTATION\DO\02-EXTRACT.do
////////////////////////////////////
* EXTRACT DATA FROM PNAD.DTA
cd $datadir
/* Legenda: SDR = semana de referência
TP = trabalho principal */
#delimit ;
use
V0101 /* ano */
UF /* unidade federativa */
V0102 /* controle */
V0103 /* série */
V0301 /* ordem */

```

```

V0403 /* n. da família */
V4611 /* peso do domicílio */
UPA /* município */
V4617 /* estrato */
V4618 /* unidade primária de amostragem */
V0302 /* sexo */
V8005 /* idade */
V0402 /* condição na família */
V0404 /* cor */
V0601 /* alfabeto */
V0602 /* estuda */
V6003 /* curso que frequenta */
V6007 /* curso mais elevado que frequentou */
V4803 /* anos de estudo */
V9001 /* trabalhou na sdr */
V9906 /* cód. de ocupacao (CBO) no TP da SDR */
V9907 /* cód. de atividade (CNAE-Dom) no TP da SDR */
V9058 /* n. de horas trabalhadas por semana no TP da SDR */
V9087 /* associado a um sindicato no mês de referência */
V9892 /* idade em que começou a trabalhar */
V4706 /* posição no TP da SDR */
V4718 /* rendimento do TP */
V4722 /* rendimento mensal familiar */
V4723 /* tipo de família */
V4724 /* n. de componentes da família */
V4728 /* região urbana ou rural */
V4107 /* região metropolitana */
using PNAD2008 , clear ;
#delimit cr
// D:\DISSERTATION\DO\03-LOG1-RAWDATA.do
////////////////////////////////////
* DESCRIPTIVE STATISTICS
gen un=1
svyset V4618 [pw=V4611], strata(V4617) singleunit(certainty)
capture log close
log using LOG1-RAWDATA , text replace
count
svy: total un
log close
// D:\DISSERTATION\DO\04-BASICVARS.do
////////////////////////////////////
* WAGE AND SELECTION EQUATION VARIABLES
*****
* WAGE EQUATION VARIABLES
*****
* Rendimento do TP
gen w = V4718 if V4718!=99999999999 & !mi(V4718)
* Ln(rendimento do TP por hora trabalhada)
gen lwph = ln(w/(V9058*4)) if !mi(w,V9058)
* Anos de estudo
gen educ = V4803-1
recode educ (16=.)
recode educ (15=17) if V6003==11 | V6007==9
* Anos de estudo ao quadrado
gen educsq = educ^2
* Anos de estudo ao cubo
gen educcb = educ^3
* Anos de experiência em dezenas de anos
gen exp = (V8005-V9892)/10 if !mi(V9892)
replace exp = (V8005-educ)/10 if (V8005-V9892)<0

```

```

replace exp=. if w==.
* Anos de experiência em dezenas de anos ao quadrado
gen expsq = exp^2
* Binária de gênero [base: masculino]
recode V0302 (2=0)(4=1), gen(fem)
lab def labfem 0 "male" 1 "female"
lab val fem labfem
* Indicativa de posição no TP da SDR [base: empregado priv. formal]
recode V4706 (1 6=0)(2=1)(3=2)(4 7=3)(9=4)(10=5)(11/13=.),gen(position)
lab def labposition 0 "empdo c/ CT" 1 "milit." 2 "func. púb." 3 "empdo s/ CT" ///
4 "conta próp." 5 "empdor"
lab val position labposition
* Indicativa de cor [base: branca]
recode V0404 (2 6 0 9=0)(4 8=1),gen(black)
lab def labblack 0 "other" 1 "black/brown"
lab val black labblack
* Binária de rural [base: região urbana]
recode V4728 (1/3=0)(4/8=1),gen(rural)
lab def labrural 0 "urban" 1 "rural"
lab val rural labrural
* Binária de pessoa de referência [base: outra]
recode V0402 (2/8=0)(1=1),gen(refpers)
lab def labrefpers 0 "other" 1 "reference person"
lab val refpers labrefpers
* Indicativa de macrorregião [base: Sudeste]
recode UF (30/39=0)(10/19=1)(20/29=2)(40/49=3)(50/59=4),gen(region)
lab def labregion 0 "Southeast" 1 "North" 2 "Northeast" 3 "south" 4 "Midwest"
lab val region labregion
* Binária de sindicato [base: não associado]
recode V9087 (3=0)(1=1),gen(union)
lab def labunion 0 "not union member" 1 "union member"
lab val union labunion
* Indicativa de ocupação [base: outra]
gen occup=0 if V9906<.
replace occup=1 if V9906>=1111 & V9906<=1230
replace occup=2 if V9906>=1310 & V9906<=1320
lab def laboccup 0 "other" 1 "chief" 2 "manager"
lab val occup laboccup
* Binária de região metropolitana [base: ã metrop.]
recode V4107 (2/3=0)(1=1),gen(metro)
lab def labmetro 0 "non-metropolitan area" 1 "metropolitan area"
lab val metro labmetro
* Binária de casado [base: ã casado]
recode V4723 (6/8=0)(1/4=1)(10=.),gen(married)
* Número de crianças <14 anos na família
gen child = V8005<14 if !mi(V8005)
bysort V0102 V0103 V0403: egen nchild = total(child)
*****
* ADDITIONAL VARIABLES FOR THE SELECTION EQUATION
*****
* Binária de trabalha [base: ã trab.]
gen works = V4718>0 & V4718<999999999999
lab def labworks 0 "does not work" 1 "works"
lab val works labworks
* Rendimento mensal familiar
recode V4722 (999999999999=.),gen(finc)
gen lfinc = ln(finc)
* Número de componentes da família
gen nfmem = V4724
* Binária de mãe solteira com criança <14 anos [base: outra]

```

```
recode V4723 (6 8=1)(10=.)(nonm=0),gen(mthchild)
* Idade em dezenas de anos
gen age10 = V8005/10
* Idade em dezenas de anos ao quadrado
gen age10sq = age10^2
// D:\DISSERTATION\DO\05-SECTORVARS.do
////////////////////////////////////
*****
* SECTOR VARIABLES
*****
* VARIABLE WITH 52 SECTORS
rename V9907 cnae5
gen cnae3 = int(cnae5/100)
tab cnae3
#delimit ;
recode cnae3
(11=1)
(12=2)
(13=3)
(14=4)
(20=5)
(50=6)
(100 110 230 234=7)
(130 140=8)
(150 160=9)
(170=10)
(180=11)
(190=12)
(200=13)
(210=14)
(220=15)
(240=16)
(250=17)
(260=18)
(270=19)
(280=20)
(290 300 310 320=21)
(330=22)
(340 350=23)
(360=24)
(370=25)
(400=26)
(410=27)
(459=28)
(500=29)
(530=30)
(531=31)
(550=32)
(600 610 620=33)
(630=34)
(640=35)
(650=36)
(660=37)
(670=38)
(700=39)
(450 710=40)
(720=41)
(730=42)
(740=43)
(750=44)
```

```

(800=45)
(850=46)
(900=47)
(910=48)
(920=49)
(930=50)
(950=51)
(990 998=52)
,gen(sector) ;
#delimit cr
tab sector
* VARIABLE WITH 11 SECTORS
#delimit ;
recode sector
(1/6=1)
(7/8 18/20=2)
(9/13 24=3)
(14/17 21/23=4)
(26/27 46=5)
(28=6)
(29/30 32 39/40 50=7)
(33/34=8)
(35/38 41/43 45 48/49=9)
(25 31 47 51/52=10)
(44=11)
,gen(mjsector) ;
label def labmjsector
1 "Agriculture"
2 "Energy & mining"
3 "Manufacturing"
4 "High tech manufact."
5 "Utilities"
6 "Building"
7 "Trade"
8 "Transport"
9 "High tech services"
10 "Other services"
11 "Government" ;
#delimit cr
label val mjsector labmjsector
tab mjsector
// D:\DISSERTATION\DO\06-HCEVARS.do
////////////////////////////////////
*****
* VARIABLES TO MEASURE HUMAN CAPITAL EXTERNALITIES
*****
* Ln da fração de pessoas com pós-graduação em cada branch [11 categ.]
bysort sector: egen nsec = count(sector)
gen educ17 = educ>=17 if !mi(educ)
by sector: egen neduc17sec = sum(educ17)
by sector: gen feduc17sec = neduc17sec/nsec
gen lfeduc17sec = ln(feduc17sec)
rename lfeduc17sec knwconc
* Fração de pessoas empregadas no setor de serviços às empresas, por UF
recode cnae5 (64020 72010 73000 74011 74012 74021 74022 74040=1) ///
(nonmis=0) , gen(bssect)
rename UF state
bysort state: egen nst = count(state)
by state: egen nbssectst = sum(bssect)
by state: gen fbssectst = nbssectst/nst

```

```

gen lfbssectst = ln(fbssectst)
rename lfbssectst knwsprd
/*
* Ln da fração de trabalhadores com nível de escolaridade k
egen tesc = count(esc)
bysort esc: egen nesc = count(esc)
by esc: gen fesc = nesc/tesc
gen lfesc = ln(fesc)
* Ln da fração de trabalhadores com grau de escolaridade k
recode esc (0/4=4)(5/8=8)(9/11=11)(12/15=15)(17=17) , gen(gesc)
bysort gesc: egen ngesc = count(gesc)
by gesc: gen fgesc = ngesc/tesc
gen lfgesc = ln(fgesc)
* Ln da fração de trabalhadores do setor j com nível de escolaridade k
bysort setor: egen tescset = count(esc)
bysort setor esc: egen nescset = count(esc)
by setor esc: gen fescset = nescset/tescset
gen lfescset = ln(fescset)
* Ln da fração de trabalhadores do setor j com grau de escolaridade k
bysort setor gesc: egen ngescset = count(gesc)
by setor gesc: gen fgescset = ngescset/tescset
gen lfescset = ln(fgescset)
* Ln da fração de pessoas com pós-graduação em cada setor [52 categ.]
bysort setor: egen nset = count(setor)
by setor: egen nesc17set = sum(esc17)
by setor: gen fesc17set = nesc17set/nset
gen lfesc17set = ln(fesc17set)
* Ln da fração de pessoas com pós-graduação em cada UF
bysort UF: egen nesc17uf = sum(esc17)
by UF: gen fesc17uf = nesc17uf/nuf
gen lfesc17uf = ln(fesc17uf)
* Média de escolaridade em cada setor (52 categorias)
bysort setor: egen mescset = mean(esc)
* Média de escolaridade dos empresários de cada setor
gen escompres = esc if pos==5
by setor: egen mescompres = mean(escompres)
* Média de escolaridade em cada branch (11 categorias)
bysort branch: egen mescbr = mean(esc)
* Média de escolaridade em cada UF (27 categorias)
bysort UF: egen mescuf = mean(esc)
* Indicadora de se a escolaridade do indivíduo é superior à média do setor
gen descsupmset = esc>mescset
* Indicadora de se a escolaridade do indivíduo é superior à média da branch
gen descsupmbr = esc>mescbr
* Fração de pessoas empregadas no setor de serviços às empresas, por UF [alt]
recode cnae5 (74011 74012 74021 74022 74030=1) (nonm=0),gen(setsae)
bysort UF: egen nsetsaeuf = sum(setsae)
by UF: gen fsetsaeuf = nsetsaeuf/nuf
gen lfsetsaeuf = ln(fsetsaeuf)
*/
// D:\DISSERTATION\DO\07-RENAME.do
////////////////////////////////////
* RENAME VARIABLES
rename V0101 year
rename V0102 control
rename V0103 series
rename V0301 order
rename V0403 idf
rename V4611 wgt
rename UPA upa

```

```

rename V4617 str
rename V4618 psu
rename V9906 cbo4
rename V9058 hours
rename V8005 age
svyset psu [pw=wgt], strata(str) singleunit(certainty)
// D:\DISSERTATION\DO\08-SINGLEPSU.do
/////////////////////////////////////////////////////////////////
* VERIFY STRATUM WITH SINGLE SAMPLING UNIT
svyset psu [pw=wgt], strata(str) singleunit(missing)
svy: heckman $T $X , select($H)
svydes if e(sample), single
* ALLOCATE SINGLE PSU STRATA INTO LARGER STRATA (5 min)
*ssc install idonepsu
set more off
foreach i of numlist 11/17 21/29 31/33 35 41/43 50/52 {
idonepsu if e(sample) & state=='i', strata(str) psu(psu) generate(new_)
replace str=new_str
replace psu=new_psu
drop new_*
}
/* Esse procedimento aloca os estratos com psu único para
estratos com maior número de psus, por UF. Porém, como este
procedimento não afeta em nada os coeficientes estimados e
nem seus desvios-padrão*/
* REVERIFY STRATUM WITH SINGLE SAMPLING UNIT
svydes if e(sample), single
// D:\DISSERTATION\DO\09-DROP.do
/////////////////////////////////////////////////////////////////
* DROP SOME OBSERVATIONS
* Who is under 20 years old
drop if age<20
* Who is over 60 years old
drop if age>60
* who works, but has zero wage
drop if w==0
* who works, but zero hours
drop if hours==0
* Who earns over R$ 30,000 per month
drop if w>30000 & !mi(w)
* If a personal information is missing that should not
drop if mi(works,finc,nfmem,age10,educ,fem,nchild,region,black,rural,refpers,metro)
* If a professional information is missing that should not
drop if mi(exp,position,union,occup,mjsector) & works==1
// D:\DISSERTATION\DO\10-LOG2-USEDATA.do
/////////////////////////////////////////////////////////////////
* DESCRIBE USED DATA
capture log close
log using LOG2-USEDATA , text replace
count
svy: total un
svy: total un, over(fem)
tabstat $cvars , stat(n mean sd min max) col(stat)
xi, noomit: tabstat $ivars , stat(sum mean) col(stat)
table age educ [pw=wgt] , c(mean w) format(%9.0f)
table mjsector educ [pw=wgt] , c(mean w) format(%9.0f)
table mjsector [pw=wgt] , c(count educ sum educ17 mean feduc17sec)
table mjsector , c(count educ sum educ17 mean feduc17sec)
table state , c(count educ sum educ17 mean fbssectst)
log close

```

```

// D:\DISSERTATION\DO\11-GRAPHS.do
/////////////////////////////////////////////////////////////////
* GRAPHS
graph bar (mean) w [pw=wgt], over(educ)
graph export "g1-bar(w)[educ].png" , replace
quietly gen exp5 = int(exp*2)
graph bar (mean) w [pw=wgt], over(educ, label(angle(90) labsz(2))) over(exp5)
graph export "g2-bar(w)[educ,exp5].png" , replace
drop exp5
graph bar (mean) w [pw=wgt], over(educ) by(mjsector)
graph export "g3-bar(w)[educ]{mjsector}.png" , replace
graph bar (mean) w [pw=wgt], over(age10) by(mjsector)
graph export "g4-bar(w)[age10]{mjsector}.png" , replace
// D:\DISSERTATION\DO\12-LOG3-RETURNS.do
/////////////////////////////////////////////////////////////////
* RETURNS TO SCALE WITH AND WITHOUT SVY AND HECKMAN
* PROGRAM FOR RETURNS TO SCALE TO EDUCATION
program define rtscale
matrix B = e(b)
scalar b1 = e1(B,1,1)
scalar b2 = e1(B,1,2)
scalar b3 = e1(B,1,3)
scalar r11 = -2*b2/(6*b3)
#delimit ;
matrix R = (
r11 , b1+2*b2*r11+3*b3*r11^2 , 0 \
4 , b1+2*b2*4+3*b3*4^2 , 2*b2+6*b3*4 \
8 , b1+2*b2*8+3*b3*8^2 , 2*b2+6*b3*8 \
11 , b1+2*b2*11+3*b3*11^2 , 2*b2+6*b3*11 \
15 , b1+2*b2*15+3*b3*15^2 , 2*b2+6*b3*15 \
17 , b1+2*b2*17+3*b3*17^2 , 2*b2+6*b3*17 )
;
#delimit cr
matrix colnames R = educ ro dro
matrix rownames R = dro0 4 8 11 15 17
end
* MAKE LOG FILE
capture log close
log using LOG3-RETURNS , text replace
/* RETURNS TO SCALE WITHOUT COMPLEX SAMPLE DESIGN (SVY)
AND HECKMAN'S (NON-SELECTION) CORRECTION BUT WITH
HETEROSCEDASTIC-ROBUST STANDARD-ERRORS */
reg $T $s $r $X , robust
estimates store ols
estimates stats
* RETURNS TO SCALE WITH SVY CORRECTION
svy: reg $T $s $r $X
estimates store svy
estimates stats
* RETURNS TO SCALE WITH SVY AND HECKMAN'S CORRECTION
svy: heckman $T $s $r $X , select($H)
estimates store svyheck
estimates stats
* SVY & HECKMAN EFFECTS ON COEFFICIENTS
estimates table ols svy , b(%9.3g) se(%9.3g) stats(N F r2)
estimates table svyheck , b(%9.3g) se(%9.3g) stats(N F rho lambda)
* SVY & HECKMAN EFFECTS ON RETURNS TO SCALE
estimates restore ols
rtscale // program
mat lis R

```

```

estimates restore svy
rtscale // program
mat lis R
estimates restore svyheck
rtscale // program
mat lis R
log close
// D:\DISSERTATION\DO\13-LOG4-POOLED.do
////////////////////////////////////
* SECTORAL RETURNS TO EDUCATION
* POOLED REGRESSION WITH SECTOR-DUMMIES
/*
#delimit ;
reg $T $s $r $X
i.mjsector#c.educ
i.mjsector#c.educsq
i.mjsector#c.educb
, robust ;
#delimit cr
*/
capture log close
log using LOG4-POOLED , text replace
#delimit ;
svy: heckman $T
i.mjsector#c.educ
i.mjsector#c.educsq
i.mjsector#c.educb
$s $r $X
, select($H) ;
#delimit cr
testparm i.mjsector#c.educ i.mjsector#c.educsq i.mjsector#c.educb
estimates table , b(%9.3g) se(%9.3g) stats(N F rho lambda)
/* If we interact $r with branch or when we include $s,
almost all branch*esc coefficients become insignificant. */
log close
// D:\DISSERTATION\DO\14-LOG5-SECTORS.do
////////////////////////////////////
* SECTORAL RETURNS TO EDUCATION (20min)
* REGRESSIONS PER SECTOR
capture log close
log using LOG5-SECTORS , text replace
*preserve
forvalues i = 1/11 {
*drop if mjsector!='i' & !mi(mjsector)
*reg $T $X $r if mjsector=='i'|mjsector==. , robust
svy, subpop(if mjsector=='i'|mjsector==.): heckman $T $r $X , select($H)
estimates store s'i'
rtscale // program
mat lis R
*restore , preserve
}
* SUMMARY OF ESTIMATES
estimates table s? s10 s11 , b(%9.3g) se(%9.3g) stats(N F rho lambda)
log close

```