fea-RP Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto
Universidade de São Paulo

Texto para Discussão

Série Economia

TD-E / 03 - 2008

O efeito do título de propriedade da terra na determinação da renda rural

Prof. Dr. Ricardo Luis Chaves Feijó

Universidade de São Paulo

Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Reitora da Universidade de São Paulo
Suely Vilela

Diretor da FEA-RP/USP
Rudinei Toneto Junior

Chefe do Departamento de Administração
Marcio Mattos Borges de Oliveira

Chefe do Departamento de Contabilidade
Maísa de Souza Ribeiro

Chefe do Departamento de Economia
Maria Christina Siqueira de Souza Campos

CONSELHO EDITORIAL

Comissão de Pesquisa da FEA-RP/USP

Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto
Avenida dos Bandeirantes,3900
14049-900 Ribeirão Preto – SP

A série TEXTO PARA DISCUSSÃO tem como objetivo divulgar: i) resultados de trabalhos em desenvolvimento na FEA-RP/USP; ii) outros trabalhos considerados de relevância dadas as linhas de pesquisa da instituição. *A série foi subdividida em função das principais áreas de atuação da FEA-RP/US: Economia, Administração e Contabilidade. Veja o site da CPq na Home Page da FEA-RP:* www.fearp.usp.br. *Informações: e-mail:* cpq@fearp.usp.br

O EFEITO DO TÍTULO DE PROPRIEDADE DA TERRA NA DETERMINAÇÃO DA RENDA RURAL

Introdução

Pretende-se neste ensaio averiguar teórica e empiricamente o papel dos títulos de propriedade da terra rural no desempenho produtivo dos estabelecimentos rurais. Acredita-se que as conclusões deste estudo podem trazer importantes implicações para as políticas públicas. Para tanto, o artigo está estruturado da seguinte forma: primeiramente apresenta-se um arrazoado conceitual e teórico apoiado na literatura. Na segunda parte, examinam-se resultados empíricos observados com o tratamento estatístico e econométrico dos dados residentes nas bases da PNAD em diferentes anos. Os dados quantitativos são apresentados e submetidos a um tratamento estatístico cujos resultados serão comentados. Em seguida, interpretam-se os resultados contrapondo-os às expectativas geradas na parte teórica. Por fim, a parte final descreve os desafios futuros da pesquisa e os possíveis desdobramentos.

Como se sabe, a cada ano dados de aproximadamente 400 mil entrevistas são disponibilizados nas PNADs. Esses dados apresentam-se em formato numérico compactado que não pode ser lido diretamente em planilhas; de forma que se faz necessária a instalação de algum programa que possa ler e tratar esses dados. Utilizamos o programa SAS para tratar microdados.[1]

Bem compreendidos o tipo de informação coletada pelas PNADs, as variáveis em questão e os aspectos metodológicos dessas variáveis, estudou-se as notas metodológicas, incluindo-se a leitura, em diferentes anos, da conceituação das características investigadas. Também se estudou o plano de amostragem, o processo de seleção e de expansão da amostra, a precisão das estimativas etc. Isto foi feito para todos os anos em que se pretendia examinar o dados das PNADs.

Este estudo parte dos achados empíricos contidos em Hoffmann e Ney (2003) e procura estender o escopo da pesquisa para outros anos, além do ano de 2001 estudado na referência, e também incorporar outras variáveis no modelo de regressão. O estudo citado concentra-se em procurar informações sobre a riqueza das famílias rurais na variável "posição na ocupação": variável que descreve se o entrevistado é empregado, trabalhador doméstico, conta própria, empregador, trabalhador não remunerado familiar, trabalhador na construção para uso próprio, trabalhador não remunerado etc. Tal estudo somente se interessou em identificar três posições: empregador, empregado e conta própria. Para tanto, ele deve ter utilizado algum critério para agrupar aquele conjunto de posições em apenas três categorias. Faremos o mesmo em relação a essa variável. Adicionalmente, veremos a importância de outra variável denominada "condição de ocupação" (variável V9021 da PNAD, "condição em relação ao empreendimento agrícola") que discriminava se o trabalhador é proprietário ou outra condição: arrendatário, posseiro etc. Muito útil essa informação para analisar o efeito da titularidade da terra rural, na medida em que essa condição poderia ser captada pela variável apontada na base da PNAD.

[1] O software SAS oferece facilidades em relação aos outros softwares econométricos, na extração e tratamento dos dados da PNAD. Os CD-ROMs de cada PNAD já vêm com o programa SAS de extração e conversão dos dados no formato de planilha. Ademais, em cada ano o CD da PNAD disponibiliza pastas e arquivos: há arquivos com o dicionário de variáveis, os arquivos de leituras com diversas informações metodológicas, e os enormes arquivos de dados que só podem ser lidos por meio de um software especial.

Outra questão técnica que se pretende examinar neste ensaio trata-se do uso de coeficientes de variação. A amostra de cerca de 400 mil observações da PNAD não é diretamente representativa da população como um todo. Há a necessidade de se utilizar o fator de expansão ou coeficiente de variação associado a cada pessoa na amostra, como fator de ponderação na hora de se obter as equações de rendimento. Há dentre as variáveis da PNAD uma que representa o peso da pessoa, ou seja, o número de pessoas que está sendo representada por uma única entrevista. Trata-se da variável V4729. Subsiste muita discussão a respeito do uso desse coeficiente de expansão na estimação das regressões com os dados da PNAD. As opiniões técnicas a respeito divergem: multiplicar todas as variáveis de uma observação pelo respectivo coeficiente de variação (peso da pessoa), ou então simplesmente desconsiderar o peso da pessoa na estimativa das regressões, dadas as distorções estatísticas produzidas pelo método anterior.

A parte teórica da pesquisa

A principal referência, sem dúvida, trata-se de Hoffman e Ney (2003). Nosso estudo continua de onde Hoffman e Ney pararam. Refizemos seus resultados, simulamos equações de rendimentos para o mesmo ano e comparamos os dois conjuntos de resultados, o nosso e o dele. É fundamental uma boa compreensão da pesquisa de Hoffmann e Ney, pois o que pretendemos nesta pesquisa é enriquecê-la incorporando novas variáveis, novas considerações e um estudo que não se limite à PNAD de um único ano (2001), como no estudo citado.

A lógica da organização econômica, a teoria da firma, a teoria dos contratos incompletos configuram temas, da coletânea editada por Williamson e Winter (1993), que nos pareceram apropriados, e mesmo essenciais, para uma compreensão teórica do papel dos direitos de propriedade, em especial da titularidade da terra, e dos seus efeitos na rentabilidade dos estabelecimentos rurais. Já a obra clássica de Willianson (1995) proporciona um estudo bastante profundo da teoria dos custos de transação, essas duas referências tratam-se de contribuições fundamentais no campo da economia institucional. Estas obras esclarecem a razão de a propriedade da terra condicionar o retorno econômico do estabelecimento rural. A teoria em questão, quando pensada para o caso em apreço, sugere o que se pode esperar das informações fornecidas pelo tratamento dos dados das PNADs, antes mesmo da obtenção dos resultados empíricos. Ou seja, a consideração teórica suscita expectativas do efeito dos direitos de propriedade, mesmo que ainda num nível bem abstrato.

O ensaio de Eggertsson, *Property rights and economic outcomes*, em Eggestsson (1999), trata-se de um dos estudos mais completos sobre o papel dos direitos de propriedade. Suas reflexões sintetizam os principais achados da disciplina que se convencionou chamar de Economia dos Diretos de Propriedade.

A Nova Economia Institucional é a base teórica dos estudos que focalizam o papel dos direitos de propriedade. Coase (1998) e Wiliamson (1995) são as referências clássicas, e também North (1981), Nelson e Winter (1982) e alguns outros. Especificamente sobre o papel dos direitos de propriedade na promoção de melhores resultados econômicos, destacam-se, nessa literatura, autores como Jensen e Mecking (1979), Hoff (2000), Pejovich (1982) e Eggestsson (1999), os quais, com base nos conceitos dos clássicos supracitados da Economia Institucional, desenvolvem toda uma linha de argumentação que explica o papel dos referidos direitos no resultado econômico.

A economia tradicional, anterior à abordagem institucional, não levava em conta a relação entre direitos de propriedade e o desempenho da unidade produtiva. Nos últimos anos, começaram a aparecer estudos em que se busca estimar como os direitos de propriedade afetam o desempenho produtivo. Na nova ótica, a análise da produção não pode mais ser feita considerando-se apenas o nível dos fatores de produção empregados e a escolha de tecnologias: deve-se considerar também o aspecto da propriedade dos recursos no cálculo da eficiência alocativa e na estimação do resultado da produção.

Este ensaio parte da premissa de que a propriedade interessa. As leis, as normas sociais e o conjunto de instituições condicionam o desempenho produtivo. Os sistemas de propriedade afetam os processos de escolha e de decisão dos agentes, e condicionam o resultado econômico. O comportamento do agente produtivo depende dos direitos de propriedade. A decisão de o quanto investir, as estratégias de produção, a escolha de técnicas e de padrões tecnológicos, a possibilidade de se estabelecer determinadas relações sociais, o monitoramento dos desvios de comportamento dos envolvidos em contratos etc., tudo isto é condicionado pelos direitos de propriedade.

A Nova Economia Institucional desenvolve diversos modelos que explicam o efeito dos direitos de propriedade no resultado da produção. Para tanto, os teóricos desta escola incorporam em suas análises conceitos tais como informação assimétrica e custos de transação que explicariam porque tais direitos contam.

Com efeito, a estrutura dos direitos de propriedade afeta o comportamento individual e a produção, pois, influenciam o leque de regras internas do jogo e afetam os custos de usar arranjos contratuais alternativos. No caso da agricultura, restrições nos direitos de propriedade reduzem o incentivo para investir em melhoria da terra no longo prazo. Considerando-se a estrutura legal dos direitos de propriedade, não se alocam recursos em projetos com retorno futuro duvidoso. Variações nos direitos de propriedade alteram as decisões alocativas dos fazendeiros e o resultado econômico.

A produção não é apenas a manipulação de processos químicos ou de engenharia; envolve, além da disposição e combinação de fatores materiais, um processo de interação humana. As possibilidades de acordos entre as pessoas dependem do arcabouço institucional, em especial, da existência ou não de direitos de propriedade.

No caso da propriedade fundiária, o produtor que detém o direito de propriedade da terra em que trabalha tem acesso a crédito e está mais propenso a atuar em cooperativas do que os não proprietários. Quem é dono da terra tem um compromisso maior com ela, investe mais, pensa no longo prazo, se preocupa mais com a conservação do solo etc. Não apenas porque pretende continuar indefinidamente a trabalhar e viver do fruto de seu empenho na terra, mas também porque a preservação do ativo fundiário em boas condições lhe possibilitará vender eventualmente a propriedade em condições favoráveis, obtendo um melhor valor por ela.

A propriedade do ativo, e da terra em particular, facilita negociações entre os atores envolvidos no processo produtivo por dois aspectos básicos: contorna problemas de informação assimétrica para os contratos (*moral hazard* e seleção adversa) e minimiza o custo de transação (confecção de contratos e cláusulas de *enforcement*).

Como os indivíduos que se envolvem em negociações detêm cada qual informação privada não revelada e efetuam ações difíceis de serem monitoradas, a eficiência desses contratos dependerá de procedimentos que funcionem como incentivos contra comportamentos oportunistas e desviantes. A economia da informação identifica diversos casos nos quais a presença do título de propriedade de ativos afeta a eficiência da produção em situações de informação imperfeita. O direito de propriedade não é um insumo de produção, na verdade ele atua de modo análogo a uma tecnologia

catalisadora da produção associada a certo uso de um vetor de insumos. Direitos de propriedade e tecnologias, portanto, agem de maneira associada. A maneira em que os direitos de propriedade facilitam as transações entre os atores envolvidos na produção é explicada nas teorias de *agente e principal*, originalmente apresentadas por Stephen Ross (1973).

Em suma, as oportunidades de trabalho dos indivíduos estão ligadas aos diretos de propriedade que eles exercem. Indivíduos com pouca propriedade (baixo padrão de riqueza material) vêem limitada a sua capacidade de realizar investimentos produtivos e de entrarem em arranjos contratuais que ensejam muito esforço. A posse de ativos mitiga os problemas de assimetria de informação; inibe os problemas de *moral hazard* e de seleção adversa ao facilitar o monitoramento pelo principal das ações dos agentes A. Tendo exposto essas várias dimensões dos direitos de propriedade, voltemos à questão principal da pesquisa que é a relação entre esses direitos e o resultado da produção.

Direitos exclusivos ou comunais afetam a produção como um fator à parte. Em algumas situações, um tipo de direito é mais eficiente, em outra, tipo diverso seria superior. Quando se alocam vários usuários em uma propriedade comum (direto comunal), pode ser que a dissipação da rentabilidade potencial do ativo comum produza um resultado inferior ao direito exclusivo. Numa análise de equilíbrio geral é difícil demonstrar qual destas modalidades de direito, ou a ausência de qualquer direito, produzirá o melhor resultado em termos de bem-estar social. A produção líquida pode até aumentar no caso de direito exclusivo, mas a perda de bem-estar dos excluídos não poderia ser comparada aos benefícios dos proprietários. O impacto depende das funções individuais de utilidade e da comparação entre elas, coisa que nem poderia ser feita. O importante é salientar que mudanças na estrutura dos direitos de propriedade produzem resultados econômicos.

Dada a presença de assimetria de informação e de custos de transação, um sistema de propriedade particular condiciona a alocação de recursos e a produção obtida. Mudanças na estrutura dos direitos de propriedade afetam o macrodesempenho de um sistema econômico. Em especial, para cada unidade produtiva, alterações na relação de posse e nos custos para exercer os direitos de propriedade afetam o resultado econômico específico.

A estrutura de propriedade altera a capacidade dos indivíduos de obter altos retornos de seu capital e de formar contratos que incorram em grandes esforços de execução. A posse de ativos afeta a produção de três maneiras que poderíamos analisar separadamente:

a) Afeta o empenho do produtor, pois ele enquanto proprietário se interessa mais pela manutenção das condições dos ativos e em estabelecer relações de longo prazo com o negócio. Por exemplo, o produtor rural que conserva a área de sua propriedade e faz investimento de longa maturação.
b) Facilita o estabelecimento de variados tipos de contratos com sócios. Tornam possíveis os pagamentos de incentivos que estimulam o esforço do sócio num cenário no qual o esforço não é visível.
c) Permite recorrer ao mercado de capital, e em especial ao crédito. A propriedade oferece uma garantia aos credores, reduz o risco de informação típico de uma operação de crédito.

Temos, portanto, base teórica para imaginar que o resultado econômico entre proprietários e não proprietários, mantidas as demais condições, seria diferente, favorecendo os primeiros. O comportamento muda quando se passa de uma condição a outra, as propensões a assumir riscos são outras, e o empenho no emprego dos fatores no processo produtivo é qualitativamente distinto. Sobre todos esses aspectos, tais

diferenças podem não ser notadas na abordagem tradicional, no entanto, no mundo real de custos de transação e informação imperfeita, é de se esperar que os retornos produtivos nos dois casos não sejam exatamente os mesmos, mesmo que se empregue a mesma combinação de fatores, na mesma tecnologia e escala.

O impacto dos direitos de propriedade na produção e na renda rural trata-se de um tema ainda recente na literatura especializada. O tipo de estudo mais freqüente focaliza os efeitos da reforma agrária na evolução da produção rural e na distribuição da renda no campo. Mas não é essa a questão que se pretende investigar neste estudo: não é o impacto produtivo da reforma agrária, mas os efeitos da titulação fundiária, mesmo que previamente o homem do campo já tenha a posse informal da terra.

Além do aspecto da qualidade dos insumos, subsiste um aspecto ainda mais crítico que é o fato de os direitos de propriedade afetarem a própria escolha de tecnologia. Como no modelo de Jensen e Meckling (1979), a firma rural depara-se com um leque de tecnologias disponíveis das quais uma delas seria adotada a depender das possibilidades contratuais. O ambiente em que se insere define as regras externas do jogo, a opção dependerá das possibilidades contratuais. Por exemplo, se o empreendedor rural, enquanto proprietário de terra, pode oferecer aos bancos colaterais como o ativo da terra, a viabilização do financiamento facilitaria o acesso a tecnologias mais avançadas de produção agrícola. Então o elemento institucional do direito de propriedade condiciona a própria função de produção; não se trata apenas de um insumo a mais, mas de uma mudança na função de produção.

Toda essa consideração de natureza teórica leva a crer que o grupo dos proprietários rurais teria *ceteris paribus* renda agrícola superior. Contudo, quando se observa uma população real composta por proprietários e não proprietários, os membros de cada um desses grupos diferem entre si em diversos aspectos a despeito de serem todos proprietários ou todos não proprietários. Agora, quando se compara aleatoriamente um membro de um grupo com membro escolhido ao acaso do outro grupo, por certo as diferenças entre eles podem ser muitas além do fato de um deles ser proprietário e o outro não.

Com efeito, o fato de o grupo de proprietário ter uma renda média superior pode estar relacionado a um conjunto de variáveis explicativas correlacionadas com o fato de ser proprietário, e não apenas ou principalmente à variável “proprietário”. Ou seja, é bem possível que os proprietários tenham maior escolaridade, mais horas mensais de trabalho, mais capital físico etc. o que explica o fato de auferirem maiores rendimentos.

Como isolar o efeito da propriedade na formação da renda? Somente com uma pesquisa de campo previamente concebida com o fito de captar esse efeito se poderia, com segurança, inferir resultados a respeito desse efeito. É bem difícil utilizar-se uma base de dados preparada para propósitos variados a fim de se isolar o referido efeito. Mesmo técnicas econométricas sofisticadas, como o *propensity score*, levam a resultados de precisão limitada, amiúde questionados. Uma forma de captar o efeito da propriedade na formação dos rendimentos seria entrevistar dois rurais com mesmas características, ou pelo menos com um amplo conjunto de características em comum. A única diferença a se considerar entre os dois entrevistados seria o fato de um deles ser proprietário e o outro não. Destarte, poder-se-ia dizer que as variações relativas na renda se devem à variável propriedade da terra.

Usando a PNAD, pensamos em identificar o efeito da variável “proprietário” de uma maneira singela por meio de equações de rendimento análogas às que foram utilizadas por Hoffmann e Ney (2003), mas com a inclusão dessa variável. Pode-se aferir o efeito da variável adicionada na qualidade da regressão observando-se a variação no R2 ajustado. O propósito, portanto, é o de estimar equações de regressão

para PNADs de diferentes anos selecionados e acompanhar inclusive a variação nos coeficientes da regressão ao longo desses anos.

Sabemos que as PNADs funcionam melhor em estudos em *cross-section* e que comparar resultados em diferentes anos tem suas limitações, tendo-se em conta, entre outros aspectos, as ligeiras mudanças anuais metodológicas e de coleta das informações. Contudo, consideramos o esforço de comparação válido, ainda que dentro de certos limites de precisão.

Então, o propósito das iniciativas em testes econométricos é o de medir o papel da variável "proprietário" na formação dos rendimentos agrícolas. Conforme o arrazoado teórico-conceitual, seria de se esperar que essa variável afetasse positivamente os rendimentos agrícolas. A inclusão dessa variável torna o modelo mais completo que o de Hoffmann e Ney (2003). Este último representa um avanço em relação a modelos anteriores que desconsideravam a contribuição de fatores associados à medida de riqueza. A posse da terra é vista pelos autores como uma medida de riqueza. Entretanto, a posse deve ser entendida como a disponibilidade de capital para o trabalho na agricultura. Significa que mais fatores estarão associados ao fator trabalho agrícola, aumentando a produtividade deste, ou então ampliando a escala do trabalho, e com isso possibilitando maiores rendimentos. Em um sentido mais estrito, a posse nem pode ser vista como riqueza, pois por riqueza deve-se sempre subentender propriedade. Aquilo de que se dispõe, mas do qual não se tem o direito de propriedade assegurado, não é rigorosamente falando representante da riqueza. Por esse conceito subtende-se algo que a pessoa possa dispor, inclusive para intercambiar, e até mesmo trocar pelo ativo mais líquido.

Entretanto, essa discussão quase semântica não afeta muita coisa no problema considerado: a posse leva a maiores rendimentos. Desconsiderar esse efeito superestima o impacto de fatores como educação na formação da renda, conforme bem apontado pelos autores supracitados. Ao lado da posse, a propriedade da terra também contribui, por sua vez e a sua maneira, na formação dos rendimentos agrícolas. A literatura já explorada nesta seção mostra as razões para tanto.

Sabe-se que a partir de 1995 o processo de reforma agrária no Brasil caminhou a passos largos. Mais de 40 mil famílias foram assentadas em 1995 sob a égide do INCRA e esse número cresceu ano a ano até um pico de 100 mil famílias-ano em 2000 e 2001. A partir de então, o número de assentados anualmente diminuiu, vindo a se recuperar em 2005. Muitos desses assentados recebem o título de proprietários da terra, de uma forma definitiva ou provisória, sendo o direito exercido de modo absoluto ou então se condiciona a titularidade da terra a uma série de obrigações da parte contemplada.

Qual o efeito deste aumento no número de proprietários rurais na determinação dos rendimentos agrícolas? A princípio, respaldado pela teoria da Economia Institucional, espera-se que a renda agrícola dos recém proprietários aumente, uma vez que agora eles podem utilizar o novo direito a fim de se beneficiarem de um número de contratos e arranjos sociais os quais não dispunham anteriormente. Mas o mundo complexo requer outras considerações se queremos dar conta do fenômeno: há que se considerar se o recém assentado já atuava como agricultor ou se era um desempregado rural, ou mesmo urbano, que pela pressão dos chamados movimentos sociais, teve reconhecido um pedaço de terra como seu, mesmo que o indivíduo nunca tenha atuado no ramo, ou que tenha sido um empregado rural sem nunca ter administrado um sítio. Pode ser que a mera propriedade da terra não resolva nada e que resulte num retumbante fracasso produtivo. E isso não apenas por falta de aptidão, mas também, convenhamos,

pela deficiência em um conjunto de iniciativas complementares de apoio ao agricultor recém-assentado: crédito, assistência técnica, facilidade de comercialização etc.

Assim, dentro desta possibilidade mais trágica, o simples fato dessa pessoa ser proprietária da terra não afeta muito a formação de rendimentos rurais. Quando se comparam os rurais como um todo, a incorporação de um novo proprietário pode:

a) Reduzir as diferenças de renda média entre os grupos de proprietários e de não proprietários.
b) Reduzir o impacto da variável "proprietário" na formação dos rendimentos agrícolas.

Se o processo de concessão de títulos da terra estiver sendo mais feliz, com os recém proprietários tendo um desempenho satisfatório, quando se comparam os rurais como um todo a incorporação de um novo proprietário pode:

a) Manter as diferenças de renda média entre os grupos de proprietários e de não proprietários (isto porque se retira um indivíduo com renda y_1 menor do grupo de não proprietários e se acrescenta o mesmo indivíduo agora no grupo de proprietários com renda y_2 maior, supondo ainda que ele gere mais renda do que a média anterior dos proprietários)[2].
b) Aumentar o impacto da variável "proprietário" na formação dos rendimentos agrícolas. O oposto da situação anterior, porque agora a variável está contribuindo para arranjos contratuais e sociais para o novo grupo de recém-assentados de uma forma até mais eficiente do que vinha contribuindo antes da expansão na titulação da terra.

A bem da verdade, dificilmente a diferença de renda entre os grupos de proprietários e de não proprietários se manterá com a concessão de novos títulos da terra. Porque os novos assentados não podem ser comparados com os antigos membros do grupo de proprietários, seguramente mais ricos, e também porque, via de regra, os recém assentados tendem a ser os mais pobres do grupo de não proprietários de onde provém. A manutenção na diferença de renda com o tempo pode se vista como um indicativo da qualidade da reforma agrária. E ainda: observar as variações dos coeficientes da variável "proprietário" é um excelente indicador da qualidade da reforma agrária, pois um aumento neste coeficiente, desde que ele seja significativo para a regressão, indica as possibilidades produtivas abertas com a titulação da terra.

Essa questão será aprofunda na análise empírica. Por ora, o que queríamos era ressaltar a importância do método de comparação entre PNADs na investigação dessas questões; justificando assim o método desta pesquisa e mostrando que ele avança em

[2] Mesmo sabendo que $y_2 > y_1$ (por hipótese), esse efeito de manter as diferenças de médias entre os grupos não necessariamente ocorrerá. A renda média do grupo de proprietários aumentará se a margem, isto é o rendimento do recém admitido membro do grupo, for maior que a média anterior do grupo. Ou seja, o novo proprietário bem sucedido terá que ter rendimento superior ao rendimento médio anterior do grupo de proprietários. Ademais, a média do grupo de não proprietários só cairá se o antigo membro, agora expelido, tivesse anteriormente um nível de rendimento maior que a média dos demais membros do grupo. Esse fato dificilmente ocorrerá, porque os principais beneficiários da reforma agrária tendem a ser agricultores bem pobres, talvez os mais pobres dentre os rurais não proprietários. Espera-se, portanto, que a média do grupo dos não proprietários aumente com a reforma agrária. Já no grupo de proprietários, os novos ingressantes teriam que ter um desempenho produtivo excepcional, de forma que sua renda fique acima da renda média anterior dos proprietários. Como os antigos proprietários de terra estão bem estabelecidos, e certamente são mais ricos, dificilmente a renda média do grupo aumentará a não ser que os novos assentados sejam excepcionalmente bem sucedidos. O caso mais provável é que mesmo com assentamentos exitosos as médias dos dois grupos irão se aproximar.

relação ao modelo mais simples e apenas aplicado ao ano de 2001 do ensaio de Hoffmann e Ney (2003).

Fechando este texto preliminar com considerações teóricas, podemos antecipar quais conclusões esperaríamos da parte empírica do trabalho, embora os dados possam desapontar:

a) Como acreditamos que os direitos de propriedade importam na formação dos rendimentos agrícolas, esperamos que a incorporação da variável "proprietário" melhore a regressão (medindo-se a qualidade dela e a intensidade da melhora pelo crescimento do R2 ajustado).
b) Que, sob o impacto da reforma agrária acentuada desde 1995, as diferenças nas média de rendimento entre os grupos de proprietário e de não proprietários estejam diminuindo.
c) Que os coeficientes de regressão da variável "proprietário" cresçam ao longo dos anos, sendo maior nas PNADs mais recentes, enquanto indicativo das novas possibilidades produtivas trazidas pela titularidade da terra, em consonância com a teoria da Nova Economia Institucional.

Parte empírica

Caracterização das PNADs e o Software SAS

A pesquisa com as entrevistas da PNAD é feita numa semana de referência; em geral, uma semana do mês de setembro. Os domicílios são visitados por milhares de pesquisadores de campo do IBGE que preenchem em cada entrevista um questionário com cerca de setenta páginas. No questionário, inicialmente o indivíduo é identificado e, em seguida, responde a perguntas agrupadas em blocos: característica da unidade domiciliar, identificação dos moradores, características gerais dos moradores, características de migração dos moradores, características de educação dos moradores, características de participação em programas sociais para educação dos moradores (5 a 17 anos de idade), características de trabalho das crianças moradoras (5 a 9 anos) e característica de trabalho e rendimento dos moradores de 10 anos ou mais de idade.[3] Em 2007 a pesquisa de campo foi feita com os entrevistadores utilizando um aparelho eletrônico *palmtop* de fácil utilização.

As entrevistas de âmbito nacional alcançaram 408.148 questionários preenchidos em 2005. Nos últimos anos, o número de entrevistados tem oscilando em torno disso. Com base na entrevista, são armazenadas, em cada qual, centenas de informações que ficam registradas em variáveis específicas. Todas as variáveis são numéricas, mas algumas são interpretadas como caracteres e devem ser convertidas, para tratamento estatístico, em variáveis numéricas. Os números das variáveis são gravados em um CD, em arquivo do tipo texto no qual as informações numéricas aparecem compactadas; isto é, o formato em que estão disponibilizados os números não é do tipo linhas e colunas (como em uma planilha). Por isso, temos que converter esses arquivos tipo texto em arquivo tipo banco de dados, em que linhas são definidas para cada indivíduo ou domicílio e colunas para as variáveis.

O próprio CD da PNAD oferece rotinas de programação que permitem fazer a conversão do arquivo texto em arquivo banco de dados. As rotinas estão em linguagem de programação SAS, mas podem ser utilizados outros programas como SPSS, S-Plus e Stata. Foi essa facilidade oferecida pelo IBGE e o fato desse Instituto estar mais

[3] Baseamo-nos no modelo do questionário de pesquisa de 2003.

familiarizado com o programa SAS que nos levou a optar por ele como ferramenta de extração e de tratamento dos dados da PNAD.

Para todos os anos considerados no estudo, estudando os códigos e o significado das variáveis da PNAD, bem como acompanhando o Dicionário de Variáveis. Constatam-se algumas alterações nas variáveis: em geral omissão de umas e inclusão de outras. Não faremos neste ensaio comentários detalhados sobre as mudanças que afetaram diretamente as informações que buscávamos extrair; contudo, foi levada em conta eventuais mudanças na definição das variáveis para diferente anos das entrevistas.

As variáveis que se iniciam por V47 são as chamadas "variáveis derivadas": não constam diretamente no questionário do entrevistador, mas são inferidas logicamente das variáveis coletadas no referido questionário. Elas foram muito importantes para a nossa pesquisa: boa parte das variáveis que extraímos das PNADs foram variáveis deste grupo. Das centenas de variáveis disponibilizadas pela PNAD, nos interessou, para efeito desta pesquisa, um conjunto delimitado de variáveis que foram basicamente aquelas mesmas selecionadas no trabalho de Hoffmann e Ney (2003), e mais a variável V9021 sobre a condição do entrevistado em relação ao empreendimento agrícola: descreve o fato de ser proprietário ou outra condição (parceiro, arrendatário etc.). A tabela 1 mostra as variáveis da PNAD que foram selecionadas no presente estudo e o significado de cada qual.

Tabela 1. Variáveis utilizadas nesta pesquisa das que constam na base da PNAD. O significado delas (para a PNAD de 2005).

Código da variável	Significado da variável
V4816	Atividade no trabalho principal: se agrícola (01) ou de outro tipo
V4719	Rendimento mensal de todos os trabalhos (pessoas > 10 anos)
V4703	Anos de estudo
V8005	Idade
V0302	Gênero
V4715	Posição na ocupação no trabalho principal: empregado, empregador e conta própria
V0404	Etnia ou cor (branco, preto, pardo, indígena, amarelo)
V9058	Horas semanais de trabalho
V0401	Condição na unidade familiar: se é arrimo (1) ou outra condição
UF	Estados da Federação
V4728	Domicílio é rural ou urbano?
V9152 v9157 v9162	Medidas de área do estabelecimento rural do empregador
V9154 v9159 v9164	Conversão da área do empregador para m^2
V9201 v9207 v9212	Medidas de área do estabelecimento rural do conta própria
V9204 v9209 v9214	Conversão da área do conta própria para m^2
V9021	Condição do entrevistado em relação ao empreendimento agrícola: proprietário=5

FONTE: Dicionário de variáveis da PNAD 2005 – arquivo de pessoas.

Além das variáveis do arquivo de pessoas, a PNAD oferece ainda o arquivo de domicílios com outras variáveis. A base de dados para domicílio é pequena, com menos de 10% do tamanho do arquivo de dados para pessoas. Não utilizaremos nenhuma variável desta base.

O SAS pode ser operado por meio de comandos de programação na linguagem dos programas escritos na janela Editor, ou então utilizando-se comandos das janelas. Utilizou-se o programa 1 para a geração da base de dados na forma de planilha que contém as variáveis que selecionamos para o presente estudo. O programa possibilitou a criação das variáveis que iremos utilizar nas regressões para a PNAD de 2005. O programa apaga algumas observações, cria variáveis nomeando pré-existentes ou com base em operação algébricas; tudo por meio de uma seqüência de comandos explicitados.

Programa 1 – Rotina de programação que gera a base de dados para a regressão na PNAD 2005.

```
libname pnad2005 'd:\Pnad2005';
data pnad2005.pes2005_2;
set pnad2005.pes2005;
/* trabalho principal agrícola: gera PNAD rural
*/
if v4816= ' '
then v4816=00;
if v4816=00 or v4816=02 or v4816=03 or
v4816=04 or v4816=05 or v4816=06 or
v4816=07 or v4816=08 or v4816=09 or
v4816=10 or v4816=11 or v4816=12 or
v4816=13
then delete;
/* variável dependente y */
if v4719=999999999999 or v4719=' ' or
v4719=00
then delete;
y=log(v4719);
/* escolaridade: variável anos de estudo v4703
*/
if v4703=01
then escolaridade=00;
if v4703=02
then escolaridade=01;
if v4703=03
then escolaridade=02;
if v4703=04
then escolaridade=03;
if v4703=05
then escolaridade=04;
if v4703=06
then escolaridade=05;
if v4703=07
then escolaridade=06;
if v4703=08
then escolaridade=07;
if v4703=09
then escolaridade=08;
if v4703=10
then escolaridade=09;
if v4703=11
/* regiões do Brasil */
if UF=11 or UF=12 or UF=13 or UF=14 or
UF=15 or UF=16 or UF=17
then NO=1;
else NO=0;
if UF=21 or UF=22 or UF=23 or UF=24 or
UF=25 or UF=26 or UF=27 or UF=28 or UF=29
then NE=1;
else NE=0;
if UF=31 or UF=32 or UF=33
then SE=1;
else SE=0;
if UF=35
then SP=1;
else SP=0;
if UF=41 or UF=42 or UF=43
then SUL=1;
else SUL=0;
if UF=50 or UF=51 or UF=52 or UF=53
then CO=1;
else CO=0;
/* domicílio é rural? */
if v4728=1 or v4728=2 or v4728=3
then domicilio=0;
else domicilio=1;
/* área do empredimento */
if v9152=99999999999
then delete;
if v9152='.'
then a=0;
else a=v9152;
if v9154='.'
then b=1;
else b=v9154;
if v9157='.'
then c=0;
else c=v9157;
if v9159='.'
then d=1;
else d=v9159;
if v9162='.'
```

```
then escolaridade=10;
if v4703=12
then escolaridade=11;
if v4703=13
then escolaridade=12;
if v4703=14
then escolaridade=13;
if v4703=15
then escolaridade=14;
if v4703=16
then escolaridade=17;
if v4703=17
then escolaridade=3;
/*idade e idade^2*/
if v8005=999
then delete;
idade=v8005/10;
idade2=idade*idade;
/*sexo*/
if v0302=2 then sexo=0;
if v0302=4 then sexo=1;
/* posições na ocupação no trabalho principal */
/*com a variável v4715 */
if v4715=' ' or v4715=14 or v4715=11
then delete;
if v4715=01 or v4715=04 or v4715=05 or
v4715=06
or v4715=07 or v4715=08
then empregado=1;
else empregado=0;
if v4715=09
then contapp=1;
else contapp=0;
if v4715=02 or v4715=03 or v4715=10 or
v4715=11 or v4715=12 or v4715=13
then empredor=1;
else empredor=0;
/* cor ou raça */
if v0404=2
then branca=1;
else branca=0;
if v0404=4
then preta=1;
else preta=0;
if v0404=6
then amarela=1;
else amarela=0;
if v0404=8
then parda=1;
else parda=0;
if v0404=0
then indigena=1;
else indigena=0;
if v0404=9
then delete;
/* horas semanais de trabalho */
If v9058=99 or v9058='.' or v9058=00
then delete;
loghtrab=log(v9058);
/* condição na unidade familiar */
then e=0;
else e=v9162;
if v9164='.'
then f=1;
else f=v9164;
g=a*b+c*d+e*f;
if empredor=1 and g=0
then delete;
if empredor=1 and g<500
then delete;
if empredor=1 and g>10000000
then delete;
if v9202=99999999999
then delete;
if v9202='.'
then h=0;
else h=v9202;
if v9204='.'
then i=1;
else i=v9204;
if v9207='.'
then j=0;
else j=v9207;
if v9209='.'
then k=1;
else k=v9209;
if v9212='.'
then l=0;
else l=v9212;
if v9214='.'
then m=1;
else m=v9214;
n=h*i+j*k+l*m;
if contapp=1 and n=0
then delete;
if contapp=1 and n<500
then delete;
if contapp=1 and n>100000000
then delete;
if empregado=1
then logarea=0;
else
if empredor=1
then logarea=log(g);
else
logarea=log(n);
if v9021=5
then proprietario=1;
else proprietario=0;
run;
```

if v0401=1
then condunidom=1;
else condunidom=0;	

Também utilizamos programas SAS que geram a regressão por mínimos quadrados de variáveis derivadas selecionadas. Utilizamos ainda uma rotina de programação SAS que gera a média das variáveis selecionadas calculadas em uma sub-base de dados geradas a partir de uma condição lógica.

Escolha de variáveis e caracterização delas

Para as equações de rendimento ajustadas pelo método dos mínimos quadrados, selecionamos um conjunto de variáveis explicativas similar à escolha feita por Hoffmann e Ney (2003). A variável dependente y é simplesmente o logaritmo neperiano do rendimento de todos os trabalhos das pessoas ocupadas. Para tanto, utilizou-se o log da variável V4719. Excluímos os casos com valor zero ou sem valor informado para essa variável, bem com as observações em que aparecia o valor "999999999999" que corresponde a sem declaração. O uso da função logarítmica é justificado por Hoffmann e Ney com o seguinte argumento:

"... pela forte assimetria da distribuição de renda e/ou pelo fato de os efeitos serem aproximadamente multiplicativos (ou proporcionais à renda)" (p. 135)

Grande parte das variáveis explicativas são binárias: assumem o valor 1 se a pessoa pertence a determinado grupo, ou zero caso não pertença. São consideradas as seguintes variáveis binárias:[4]

a) **Sexo**: variável binária para sexo. Assume 1 para mulheres.
b) **Contapp** e **Empredor**: duas variáveis binárias para distinguir três posições na ocupação no trabalho principal. Assumem respectivamente o valor 1 para conta própria e empregador; a condição não representada, por exclusão, é a de empregado (tomado como base). Utilizou-se para caracterizá-las a variável da PNAD V4715. Foi adotada a seguinte associação: quando a variável assume os valores em branco (sem registro), 14 (sem declaração) e 11 (não remunerado) a observação é excluída. Quando o valor dela é 1 (empregado com certeira), 4 (outros empregados sem carteira), 5 (empregados sem declaração de carteira), 6 (trabalhador doméstico com carteira), 7 (trabalhador doméstico sem carteira), 8 (trabalhador doméstico sem declaração de carteira), então considera-se a pessoa empregada, não afetando as variáveis binárias em questão. Quando a V4715 assume o valor 9 a pessoa é conta própria (1 para Contapp e 0 a ela caso contrário). Quando a V4715 assume os valores 2 (militar), 3 (funcionário público estatutário), 10 (empregador), 12 (trabalhador na produção para o próprio consumo), e 13 (trabalhador na construção para o próprio uso) a pessoa é empregadora (1 para Empredor e 0 a ela caso contrário).[5]

[4] A descrição se aplica a todas as PNADs analisadas, com ligeira mudança na variável original da PNAD utilizada em um ou outro caso. Eventuais mudanças ficarão evidentes na leitura das rotinas de programação que serão explicitadas mais adiante.

[5] Não consideramos apenas o caso 10 "empregador" como estando na categoria expressa na variável binária Empredor (empregador). Pensamos assim: um militar, ou funcionário público ou trabalhador para o próprio consumo e uso deve estar forçosamente contratando alguém para ajudá-lo na empreitada agrícola; eles não teriam tempo ou condições de atuarem por si sós na atividade agrícola, são portanto empregadores.

c) **Preta**, **parda**, **indígena** e **amarela**: quatro binárias que assumem cada qual valor 1 para distinguir cor (etnia). Branca não representada é a cor básica. Usa-se, para tanto, a variável V0404, excluindo-se as observações sem declaração de cor.
d) **Condunidom**: variável binária que diferencia a condição da pessoa na unidade familiar. Usamos a variável PNAD V0401. Se for a pessoa de referência (arrimo), a binária assume o valor 1 e zero caso contrário (cônjuge, filho, pensionista, empregado doméstico etc.)
e) **Proprietario**: essa variável binária não foi utilizada por Hoffmann e Ney (2003). Trata-se de um importante acréscimo da pesquisa na medida em que tentamos captar o efeito dos direitos de propriedade na equação dos rendimentos. Utilizamos a variável PNAD V9021 (condição em relação ao empreendimento agrícola). Quando esta vale 5 (proprietário) a variável "proprietário" assume o valor unitário e zero caso contrário (V9021 assume os valores 1 (parceiro), 2 (arrendatário), 3 (posseiro), 4 (cessionário), 6 (outra condição), 9 (sem declaração) e não aplicável).

Além dessas binárias, incorporamos no modelo algumas variáveis numéricas. São elas:

a) **Escolaridade**: anos de estudo expresso na variável V4703 com a devida interpretação do significado dos tipos numéricos. Quando assume o número 16, significando 15 ou mais anos de estudo, a variável "escolaridade" assume o valor 17, e quando a V4703 é 17 (não determinados e sem declaração) assumimos que a escolaridade destas observações é hipoteticamente de três anos de estudos (isto para não se excluir desnecessariamente algumas observações, três é um tempo representativo embora isso gere certa imprecisão na descrição da realidade).
b) **Idade**: a idade da pessoa medida em dezenas de anos. A idade é medida em dezenas de anos para evitar coeficientes muito pequenos. Usa-se a variável V8008, excluindo-se as observações sem declaração. Naturalmente divide-se seu valor por 10.
c) **Idade2**: o quadrado da variável "idade", para captar a hipótese de que a renda não varie linearmente com a idade.
d) **Loghtrab**: o logaritmo do número de horas semanais de trabalho. Segundo Hoffmann e Ney (2003): *"o coeficiente dessa variável é a elasticidade do rendimento em relação ao tempo semanal de trabalho"* (p. 136). Usa-se a variável PNAD V9058 (o seu log), excluindo-se as observações sem registro, sem dados ou com valor zero.
e) **Logarea**: uma variável destinada a captar o efeito da área do empreendimento. Toma-se o logaritmo da área, obtido utilizando-se um cálculo sobre diversas variáveis da PNAD conforme a parte intitulada "área do empreendimento" do programa 1. As variáveis de V9152 a V9164 referem-se a empregadores e as de V9202 a V9214 diz respeito aos de conta própria. V9152, V9157, V9162 são as áreas de três parcelas de terra do empreendimento. V9154, V9159, V9164 são as equivalências das áreas em m^2 quando se multiplica a área pelo valor da equivalência. A soma dos três produtos fornece a área total do empregador. Note que eliminamos as observações sem área declarada e que quando não havia informação de equivalência em m^2 a variável para equivalência assume o valor 1. Todos os empregadores com área nula foram eliminados da amostra. E também os empregadores com área abaixo de 500 m^2 e acima de 100 milhões de m^2 seguindo sugestão de Hoffmann e Ney (2003): "*As informações sobre a área do empreendimento, por exemplo, não abrangem todas as atividades agrícolas, há dados de empreendimentos com área demasiadamente grande e também existem*

dados estranhos na cauda inferior da distribuição da terra. Optamos por excluir os empregadores e de conta própria que não têm área declarada na PNAD ou declararam área menor ou igual a 0,05 ha ou área maior ou igual a 10.000 ha." (p.116) V9202, V9207, V9209 são as áreas de três parcelas de terra do empreendimento para os de conta própria. V9104, V9109, V9112 são as equivalências, para esse grupo, das áreas em m^2 quando se multiplica a área pelo valor da equivalência. A soma dos três produtos fornece a área total do conta própria. Note que também eliminamos as observações do conta própria sem área declarada e que quando não há informação de equivalência em m^2 a variável assume o valor 1. Todos os conta própria com área nula foram eliminados da amostra. E também os conta própria com área abaixo de 500 m^2 e acima de 100 milhões de m^2. Em todos os empregados a variável logarea assume o valor zero. Caso seja empregador ou conta própria tomamos o logaritmo da respectiva área calculada anteriormente.

Em acréscimo à definição das variáveis que seriam empregadas no modelo, tivemos inicialmente que preparar a base de dados para a PNAD rural, que não se confunde com a PNAD como um todo. Tínhamos que definir um critério de extração da parte rural da PNAD. Adotamos então a variável V4708, de 1992 e 1993, renomeada, porém com o mesmo significado e descrição, V4716 na PNAD de 1999, V4766, em 2001, e V4816 em 2002, 2003 e 2005. Ela significa os "grupamentos de atividade no trabalho principal do período de referência de 365 dias para pessoas de 10 anos ou mais de idade". Eliminamos todas as observações nas quais a atividade não fosse declaradamente agrícola.[6] Com isso, o tamanho da PNAD ficou consideravelmente menor. Para o ano de 2005, enquanto a PNAD completa tem 408.148 observações, a PNAD rural apresenta 18547 observações: menos de 5% da amostra completa.

Resultados empíricos observados com o tratamento estatístico e econométrico dos dados residentes nas bases da PNAD em diferentes anos

Resultados para a PNAD de 2001 e comparação com os resultados de Hofmann e Ney

Antes de apresentarmos a análise comparativa entre as PNADs de 1992, 1993, 1999, 2001, 2003, 2003 e 2005, vamos aplicar a metodologia desta pesquisa na PNAD de 2001 com o fito de estabelecer uma comparação detalhada de nossos resultados com os obtidos por Hoffmann e Ney (2003) para o mesmo ano. Inicialmente é necessário adiantar algumas observações que nos levam a antecipar que não seria de se esperar os mesmos resultados nos dois estudos, mesmo em se tratando da mesma PNAD, para o mesmo ano. Há alguns pontos não muito claros no trabalho que estamos comparando. Primeiramente, não fica explícita a maneira como eles geraram a PNAD rural. Qual variável foi utilizada como controle e qual o critério de seleção de observações? Depois, não se conhece como foi feita a alegada ponderação na estimativa das equações de rendimentos por mínimos quadrados. Não se sabe exatamente como os autores usaram o fator de expansão associado a cada pessoa da amostra como fator de ponderação; tema sobre o qual discorreremos com profundidade mais adiante. Viceja também certa obscuridade em relação à variável binária Z (sic) que assume 1 se o indivíduo é empregador ou conta própria e zero quando é empregado (p. 136). Por que assumir essa

[6] Não há a categoria de atividade pecuária ou outras que se possam associar a atividades rurais não agrícolas.

nova variável e não simplesmente apagar as observações quando empregador e conta própria tiverem área zero e assumir que todos os empregados devem possuir área rural zero, conforme feito por nós?

Hoffmann e Ney estimam quatro equações de rendimento: no modelo 1, ele não incorpora as variáveis "posição na ocupação" e "área". No modelo 2, inclui a "posição na ocupação", no modelo 3 inclui ambas: "posição na ocupação" e "área", o modelo 4 trata "escolaridade" com 15 variáveis binárias para os coeficientes da regressão para cada ano de estudo. O objetivo do estudo supracitado é o de mostrar como o coeficiente da regressão para escolaridade é afetado pela incorporação de variáveis que mensuram, de alguma forma, o grau de riqueza do indivíduo, quer pela "posição na ocupação", que está correlacionada com posse da terra, ou pelo tamanho da área declarada.

Nossa preocupação é outra: a de como as próprias estimativas dos autores são afetadas pela inclusão de uma nova variável indicadora da propriedade (não apenas posse) da terra. Tomando-se o modelo 3, que inclui "área" e "posição na ocupação", o propósito desta seção é o de comprar os resultados desse outro estudo com o nosso a respeito dos coeficientes das variáveis na regressão e a "renda esperada" para as variáveis binárias (diferença percentual entre os rendimentos estimados na categoria considerada e na categoria tomada como base).

Fazendo-se a comparação entre os dois trabalhos, pode-se observar na tabela 2 que as duas PNADs rurais têm quase o mesmo tamanho: pouco acima de 16 mil observações; o que demonstra que ambos utilizaram aproximadamente o mesmo critério de seleção da amostra. Aceitando-se que as diferenças entre os resultados não sejam muito importantes caso as discrepâncias não ultrapassem 30% do valor estimado por Hoffmann e Ney (2003), então concluímos que os resultados são semelhantes para intercepto, escolaridade, idade, idade2, cor (exceto indígena), loghtrab, condunidom, regiões do país, domicilio e logarea. Ademais, os resultados são quase idênticos (< 10% de diferença) para os coeficientes das variáveis escolaridade (diferença de 3,7%), parda (3,81%), condunidom (-3,12%), regiões do país (entre -1,14% e 8,6%), domicílio (-4,52%). Para as duas variáveis principais que estão sendo analisadas – escolaridade e logarea –, as diferenças entre nossas estimativas e a de Hoffman e Ney não são muito grandes (3,7% e -12,8% respectivamente).

Tabela 2 - Comparação entre os resultados da presente pesquisa e os resultados de Hoffmann e Ney (2003)

Variável	Coeficientes	Coeficientes em Hoffmann e Ney(2003)[2]	Diferença	Renda esperada[1]	Renda esperada em Hoffmann e Ney (2003)[2]	Diferença
intercepto	2,61746	2,1498	21,75%			
escolaridade	0,06253	0,0603	3,70%			
idade	0,36099	0,314	14,96%			
idade2	-0,03867	-0,0326	18,62%			
sexo	-0,22897	-0,16	43,11%	-36,39	-14,9	144,23%
contapp	-1,8152	-0,3702	390,33%	-5,07	-30,9	-83,59%
empredor	-1,06931	0,3294	-224,62%	250,1	39	541,28%
preta	-0,10037	-0,0841	19,35%	-35,8	-8,1	341,98%
indigena	-0,12873	0,0632	-303,69%	-40,02	6,5	-515,69%
Amarelo	0,61503	0,5542	10,98%	171,41	74,1	131,32%
parda	-0,10755	-0,1036	3,81%	-35,92	-9,8	266,53%
loghtrab	0,37372	0,5125	-27,08%			
condunidom	0,19328	0,1995	-3,12%	44,88	22,1	103,08%
NO	0,23421	0,2369	-1,14%	77,27	26,7	189,40%
SE	0,21046	0,1938	8,60%	61,46	21,4	187,20%

SP	0,56154	0,5562	0,96%	157,54	74,4	111,75%
SUL	0,34698	0,3208	8,16%	109,8	37,8	190,48%
CO	0,48161	0,4666	3,22%	142,07	59,5	138,77%
domicilio	-0,09624	-0,1008	-4,52%	-0,2	-9,6	-97,92%
logarea	0,15936	0,1815	-12,20%			
no obser	16446	16221	1,39%			
R2	0,3872	0,416	-6,92%			
R2 ajustado	0,3865	*				

FONTE: Elaboração do autor.

[1] Diferença percentual entre os rendimentos estimados na categoria considerada e na categoria tomada como base

[2] Hoffmann e Ney (2003), p. 138-139

As nossas estimativas para os coeficientes de regressão foram geradas na regressão por mínimos quadrados ordinários aplicada na base de dados obtida por um programa similar ao programa 1, mas adaptado às peculiaridade da PNAD de 2001. Infelizmente não sabemos, até o momento, qual o programa (provavelmente também escrito em SAS) que Hoffman e Ney utilizaram para gerar a base de dados utilizada em suas estimativas. Mas as diferenças nos resultados não chamam tanto a atenção exceto para as variáveis sobre "condição do produtor" (contapp (diferença de 390,33%) e empredor (-224,62%)), e para a binária indigena (-303,59%). Causa-nos estranheza o fato de os atores terem obtido um coeficiente positivo para indígena, quando sabemos a situação de penúria em que se encontra esse segmento da população. Todas as nossas simulações em todas as PNADs resultaram num desempenho muito franco para essa etnia, tanto na estimativa dos coeficientes de regressão quanto na "renda esperada". Com relação à condição do produtor, de fato, é curioso que em nossa estimativa o coeficiente de empregador (variável "empredor") apareça ligeiramente negativo, indicando que, na estimação da renda agrícola, a condição de empregador tenha efeito negativo no resultado final. No entanto, note que a renda esperada é 250,1% maior nesse segmento do que nas pessoas que são trabalhadores (quinta coluna da tabela). Os de conta própria são ligeiramente mais pobres do que os empregados (-5,07%) e seria de se esperar que o coeficiente de determinação fosse menor, em módulo, como no caso do outro estudo.

Contudo, esses coeficientes para as duas variáveis binárias de "posição na ocupação" saíram-se bem semelhantes em todas as PNADs analisadas. Acreditamos que nossa estimativa é correta e que algum detalhe possa ter influenciado os resultados de Hoffmann e Ney (2003) neste tocante. Uma das possibilidades, é a maneira como os autores fizeram a expansão da amostra usando o fator de expansão associado a cada pessoa da amostra como fator de ponderação. Tentamos um exercício de expansão da amostra simplesmente multiplicando cada observação (os dois membros da equação do modelo de regressão são multiplicados) pelo peso da pessoa. Os resultados praticamente não se alteraram conforme mostraremos adiante. Em suma, permanecemos sem saber o que afastou nossa estimativa da que foi feita pelos autores citados; uma vez que utilizamos a mesma base de dados e o mesmo modelo geral de regressão.

Ainda mais significativas foram as diferenças, para variáveis binárias, entre as "rendas esperadas" que estimamos e aquelas obtidas pelo outro estudo. As nossas diferenças superam 100% para quase todas as variáveis binárias exceto contapp (-83,59%) e domicílio (-97,92%). Não sabemos como os autores do outro estudo calcularam a "renda esperada", gostaríamos de deixar explicitado nosso método de cálculo.

Para cada variável binária, primeiramente geramos uma base de dados em que ocorre o evento tomado como base (a ocorrência em que não se atribui variável e que

funciona como base de comparação). Estimamos para essas observações a renda esperada. Na medida em que a variável de renda em nosso modelo de regressão é o logaritmo da renda, consideramos o algarismo neperiano *e* elevado ao resultado da renda estimada, que na verdade é o log da renda. Portanto, $e^{logy} = y$. A renda esperada (o log dela) é estimada multiplicando-se o valor médio de cada variável, média calculada na base de dado em questão, pelo coeficiente estimado da respectiva variável gerado no modelo de regressão aplicado na mesma base. Assim, por exemplo, para a variável cor, estima-se a renda dos brancos (categoria base) e compara-se com a renda estimada em cada categoria: pretos, pardos, indígenas e amarelos. Calcula-se então a diferença percentual (Tabela 3).

Tabela 3 – Renda esperada das variáveis binárias na PNAD de 2001

Variáveis	log y	y	Renda esperada[1]
branco	5,4040	222,30	*
preto	4,9609	142,72	-35,8%
pardo	4,9589	142,44	-35,9%
indígena	4,8929	133,34	-40,0%
amarelo	6,4027	603,49	171,5%
homem	5,1833	178,27	*
mulher	4,7310	113,40	-36,4%
não arrimo	4,8727	130,67	*
arrimo	5,2434	189,31	44,9%
NE	4,7546	116,12	*
NO	5,3271	205,84	77,3%
SE	5,2337	187,48	61,5%
SP	5,7006	299,06	157,5%
SUL	5,4956	243,62	109,8%
CO	5,6387	281,09	142,1%
empregado	5,0960	163,37	*
conta própria	5,0440	155,10	-5,1%
empregador	6,3491	571,98	250,1%
não rural	5,1520	172,78	*
rural	5,1540	173,12	-0,2%
não proprietário	5,0292	152,81	*
proprietário	5,3851	218,12	42,7%

[1] Diferença percentual entre os rendimentos estimados na categoria considerada e na categoria tomada como base

* Categoria tomada como base.

Análise comparativas entre sete PNADs: 1992, 1993, 1999, 2001, 2002, 2003 e 2005

Feita essa comparação inicial para a PNAD de 2001, estimativas dos coeficientes das variáveis e de renda esperada, entre este estudo e o de Hoffman e Ney (2003), iremos agora comparar os coeficientes de regressão para todas as variáveis binárias nas estimativas do rendimento rural nas PNADs de 1992, 1993, 1999, 2001, 2002, 2003 e 2005. A escolha destes anos não foi aleatória, na verdade tentou-se captar momentos importantes da economia rural ligados ao calendário político (início do

governo de Fernando Henrique Cardoso; início do segundo mandato; início do governo atual) e à disponibilidade das PNADs. A PNAD de 2001 foi investigada para se confrontar com os resultados de Hoffmann e Ney. A PNAD de 2003 foi escolhida aleatoriamente para estender a série examinada.

Primeiramente, rodamos as rotinas de programação do SAS que geram as bases de dados para a PNAD rural, já com as novas variáveis criadas. Por economia de espaço, não iremos apresentar os respectivos programas SAS que geram cada uma das PNADs rurais desses anos considerados. Para cada base de dados rodamos um modelo de regressão acionando rotinas de programação específicas que seguem o modelo do programa 1. Nessas regressões, consideramos de início apenas as mesmas variáveis tratadas por Hoffmann e Ney (2003). Os coeficientes estimados em cada ano aparecem na tabela 3. Em segundo passo, incorporamos a variável "proprietario", estimamos então uma nova equação de rendimentos para pessoas ocupadas na agricultura, a qual difere pela inclusão da condição do indivíduo, se proprietário ou não. O efeito de variáveis como escolaridade e área é captado agora de outra forma, pois estamos considerando uma *proxy* para propriedade da terra, diferente do modelo de Hoffman e Ney (2003) que só se preocupara em incorporar *proxies*, tais como "posição na ocupação" e área, que captam capital físico. A tabela 4 apresenta os coeficientes de regressão com e sem a incorporação da variável "proprietario". Ela ainda informa o valor da diferença percentual entre nossas estimativas do coeficiente de regressão, em diferentes anos, e as estimativas desse coeficiente para a PNAD de 2001 feita por Hoffmann e Ney. A tabela 4 informa ainda as diferenças percentuais nos coeficientes de regressão quando se inclui a variável propriedade (última coluna). Note que também consta o número de observações da PNAD rural, o R2 e o R2 ajustado.

Tabela 4 – Equações de rendimento estimadas para pessoas ocupadas na agricultura brasileira.
Coeficientes de regressão com e sem a incorporação da variável "proprietario". PNAD: 1992, 1993, 1999, 2001, 2002, 2003 e 2005.

Variável	Coeficientes	Diferenças nas estimativas em relação a Hoffmann e Ney (2003)	Coeficientes com a inclusão da variável proprietário	Diferenças nas estimativas em relação a Hoffmann e Ney (2003)	Variação nos coeficientes com a variável proprietário
PNAD 1992					
intercepto	10,2259	*375,67%*	10,2114	*374,99%*	
escolaridade	0,0697	15,56%	0,0688	14,15%	-1,22%
idade	0,3612	15,04%	0,3622	15,36%	0,28%
idade2	-0,0387	18,80%	-0,0397	21,87%	2,58%
sexo	-0,2295	42,12%	-0,2342	45,03%	2,04%
contapp	-1,3397	*-261,89%*	-1,3189	*-256,26%*	-1,56%
empredor	-0,6032	*-283,11%*	-0,6031	*-283,08%*	-0,01%
preta	-0,0850	1,12%	-0,0739	-12,13%	-13,10%
indigena	-0,3979	*-729,56%*	-0,3734	*-690,81%*	-6,16%
parda	-0,0781	-24,61%	-0,0714	-31,10%	-8,60%
amarela	0,4402	-20,58%	0,4507	-49,72%	2,40%
loghtrab	0,4228	-17,50%	0,4286	-16,38%	1,36%
condunidom	0,1952	-2,15%	0,1936	-2,96%	-0,83%
NO	0,1637	-30,88%	0,1704	-28,08%	4,06%
SE	0,2219	14,52%	0,2285	17,92%	2,97%

SP	0,5906	6,19%	0,5962	7,18%	0,94%
SUL	0,3433	7,02%	0,3437	7,14%	0,11%
CO	0,3428	-26,53%	0,3510	-24,77%	2,40%
domicilio	-0,1131	12,21%	-0,1193	18,38%	5,50%
logarea	0,1293	-28,75%	0,1165	-35,81%	-9,91%
proprietario			0,2165		
no obser	17110		17110		
R2	0,3781		0,3825		
R2 ajustado	0,3774		0,3818		
		PNAD 1993			
intercepto	6,5904	*206,56%*	6,5953	*206,79%*	0,07%
escolaridade	0,0679	12,60%	0,0661	9,59%	-2,68%
idade	0,0200	*-93,64%*	0,0101	*-96,80%*	-49,70%
idade2	-0,0003	*-99,15%*	-0,0002	*-99,46%*	-36,96%
sexo	-0,2199	36,15%	-0,2321	43,68%	5,53%
contapp	-1,4743	*298,23%*	-1,3905	*275,61%*	-5,68%
empredor	-0,7679	*-133,12%*	-0,7060	*-114,32%*	-8,06%
preta	-0,0966	14,89%	-0,0852	1,33%	-11,80%
indigena	-0,3674	*-681,30%*	-0,3283	*-619,48%*	-10,63%
parda	-0,1115	7,66%	-0,1012	-2,36%	-9,32%
amarela	0,1632	*-70,55%*	0,1445	*-83,88%*	-11,45%
loghtrab	0,4685	-8,58%	0,4760	-7,13%	1,58%
condunidom	0,3675	84,22%	0,3664	83,64%	-0,31%
NO	0,4010	69,28%	0,4032	70,21%	0,55%
SE	0,2250	16,07%	0,2306	19,00%	2,52%
SP	0,6755	21,44%	0,6817	22,57%	0,93%
SUL	0,5338	66,41%	0,5329	66,13%	-0,17%
CO	0,5348	14,61%	0,5447	16,73%	1,85%
domicilio	-0,0802	-20,45%	-0,0916	-9,17%	14,18%
logarea	0,1538	-15,24%	0,1316	-27,49%	-14,46%
proprietario			0,2900		
no obser	16554		16554		
R2	0,4247		0,4315		
R2 ajustado	0,4236		0,4305		
		PNAD 1999			
intecepto	2,7490	27,87%	2,7517	28,00%	0,10%
escolaridade	0,0604	0,12%	0,0593	-1,63%	-1,74%
idade	0,3460	10,18%	0,3454	10,00%	-0,16%
idade2	-0,0390	19,48%	-0,0396	21,44%	1,64%
sexo	-0,2312	43,17%	-0,2381	47,43%	2,98%
contapp	-1,5406	*316,16%*	-1,4771	*299,00%*	-4,12%
empredor	-0,7314	*-122,05%*	-0,6683	*-102,87%*	-8,64%
preta	-0,1232	46,49%	-0,1146	36,27%	-6,98%
indigena	-0,3149	*-598,21%*	-0,3449	*-645,66%*	9,52%
parda	-0,1306	26,08%	-0,1252	20,81%	-4,18%
amarelo	0,3398	-38,69%	0,3295	-40,55%	-3,04%

loghtrab	0,3474	-32,21%	0,3505	-31,60%	0,90%
condunidom	0,1822	-8,67%	0,1803	-9,64%	-1,06%
NO	0,1741	-26,51%	0,1622	-31,55%	-6,85%
SE	0,2065	6,56%	0,2063	6,44%	-0,11%
SP	0,5544	-0,32%	0,5595	0,60%	0,92%
SUL	0,2034	-36,59%	0,1980	-38,28%	-2,67%
CO	0,3430	-26,50%	0,3482	-25,37%	1,53%
domicilio	-0,1056	4,72%	-0,1108	9,87%	4,92%
logarea	0,1422	-21,64%	0,1264	-30,34%	-11,10%
proprietario			0,1881		
no obser	16288		16288		
R2	0,4145		0,4187		
R2 ajustado	0,4138		0,4180		
		PNAD 2001			
Intercepto	2,6175	21,75%	2,6339	22,52%	0,63%
escolaridade	0,0625	3,70%	0,0615	2,01%	-1,63%
idade	0,3610	14,96%	0,3587	14,23%	-0,64%
idade2	-0,0387	18,62%	-0,0392	20,09%	1,24%
sexo	-0,2290	41,78%	-0,2371	46,82%	3,56%
contapp	-1,8152	*390,33%*	-1,7679	*377,56%*	-2,60%
empredor	-1,0693	*-224,62%*	-1,0307	*-212,91%*	-3,61%
preta	-0,1004	19,35%	-0,0967	14,98%	-3,66%
indigena	-0,1287	*-303,69%*	-0,1176	*-286,12%*	-8,62%
Amarelo	0,6150	10,98%	0,6167	11,28%	0,27%
parda	-0,1076	3,81%	-0,1040	0,38%	-3,31%
loghtrab	0,3737	-27,08%	0,3749	-26,85%	0,31%
condunidom	0,1933	-3,12%	0,1952	-2,17%	0,98%
NO	0,2342	-1,14%	0,2282	-3,69%	-2,59%
SE	0,2105	8,60%	0,2079	7,29%	-1,20%
SP	0,5615	0,96%	0,5620	1,03%	0,07%
SUL	0,3470	8,16%	0,3390	5,66%	-2,31%
CO	0,4816	3,22%	0,4777	2,37%	-0,82%
domicilio	-0,0962	-4,52%	-0,1032	2,40%	7,25%
logarea	0,1594	-12,20%	0,1446	-20,35%	-9,28%
proprietario			0,1933		
no obser	16446		16446		
R2	0,3872		0,3909		
R2 ajustado	0,3865		0,3901		
		PNAD 2002			
Intercepto	2,5601	19,09%	2,5618	19,16%	0,07%
escolaridade	0,0592	-1,76%	0,0583	-3,32%	-1,59%
idade	0,3388	7,91%	0,3408	8,54%	0,58%
idade2	-0,0366	12,24%	-0,0375	15,06%	2,51%
sexo	-0,1864	15,40%	-0,1947	20,56%	4,47%
contapp	-1,7640	*376,51%*	-1,7225	*365,30%*	-2,35%
empredor	-0,7726	*-134,55%*	-0,7344	*-122,94%*	-4,95%

preta	-0,0571	-32,09%	-0,0530	-36,94%	-7,14%
indigena	-0,2583	*-508,66%*	-0,2780	*-539,92%*	7,65%
amarela	0,3066	-44,67%	0,2832	-48,89%	-7,63%
parda	-0,0782	-24,56%	-0,0738	-28,80%	-5,63%
loghtrab	0,4221	-17,63%	0,4250	-17,07%	0,68%
condunidom	0,2118	6,19%	0,2090	4,74%	-1,36%
NO	0,1665	-29,74%	0,1636	-30,95%	-1,73%
SE	0,2397	23,68%	0,2378	22,68%	-0,81%
SP	0,5390	-3,09%	0,5385	-3,18%	-0,10%
SUL	0,4380	36,53%	0,4319	34,64%	-1,38%
CO	0,4480	-3,99%	0,4482	-3,95%	0,04%
domicilio	-0,0966	-4,15%	-0,1032	2,40%	6,83%
logarea	0,1549	-14,68%	0,1415	-22,03%	-8,61%
proprietario			0,1792		
no obser	16671		16671		
R2	0,4187		0,4220		
R2 ajustado	0,4180		0,4213		
		PNAD 2003			
Intercepto	2,8589	32,99%	2,8733	33,65%	0,50%
escolaridade	0,0546	-9,45%	0,0536	-11,09%	-1,81%
idade	0,0272	-91,35%	0,0198	-93,70%	-27,15%
idade2	-0,0003	-99,22%	-0,0002	-99,44%	-28,06%
sexo	-0,2058	27,44%	-0,2104	30,30%	2,24%
contapp	-1,7017	*359,68%*	-1,6584	*347,97%*	-2,55%
empredor	-0,9076	*-175,52%*	-0,8769	*-166,21%*	-3,38%
preta	-0,1430	70,04%	-0,1381	64,26%	-3,40%
indigena	-0,3267	*-616,87%*	-0,3142	*-597,14%*	-3,82%
parda	-0,1030	-0,57%	-0,0969	-6,44%	-5,90%
amarela	0,3343	-39,67%	0,3171	-42,78%	-5,14%
loghtrab	0,5206	1,58%	0,5252	2,48%	0,89%
condunidom	0,2484	24,53%	0,2494	25,00%	0,38%
NO	0,2799	18,13%	0,2711	14,44%	-3,13%
SE	0,2758	42,33%	0,2706	39,62%	-1,91%
SP	0,6170	10,94%	0,6172	10,97%	0,03%
SUL	0,5112	59,35%	0,5004	55,98%	-2,12%
CO	0,4985	6,83%	0,4959	6,29%	-0,51%
domicilio	-0,0478	-52,62%	-0,0579	-42,52%	21,31%
logarea	0,1483	-18,30%	0,1343	-26,01%	-9,43%
proprietario			0,1917		
no obser	16910		16910		
R2	0,4031		0,4067		
R2 ajustado	0,4024		0,4060		
		PNAD 2005			
intercepto	2,5746	19,76%	2,5939	20,66%	0,75%
escolaridade	0,0588	-2,47%	0,0577	-4,28%	-1,85%
idade	0,3505	11,63%	0,3476	10,71%	-0,83%

idade2	-0,0352	8,01%	-0,0358	9,85%	1,70%
sexo	-0,2022	25,20%	-0,2110	30,66%	4,36%
contapp	-1,8384	*396,59%*	-1,7832	*381,69%*	-3,00%
empredor	-1,0257	*-411,38%*	-0,9677	*-193,76%*	-5,66%
preta	-0,0321	-61,85%	-0,0294	-65,04%	*-8,35%*
indígena	-0,2658	*-520,57%*	-0,2635	*-516,93%*	-0,87%
parda	-0,0505	-51,30%	-0,0451	-56,47%	-10,60%
amarela	0,1250	-77,44%	0,1000	-81,96%	*-20,02%*
loghtrab	0,4709	-8,13%	0,4734	-7,63%	0,54%
condunidom	0,1906	-4,49%	0,1891	-5,22%	-0,77%
NO	0,4212	77,78%	0,4049	70,93%	-3,85%
SE	0,2965	53,01%	0,2933	51,33%	-1,10%
SP	0,6574	18,19%	0,6572	18,15%	-0,04%
SUL	0,5121	59,64%	0,5059	57,69%	-1,22%
CO	0,5481	17,47%	0,5466	17,13%	-0,28%
domicilio	-0,0193	-80,85%	-0,0290	-71,26%	*50,10%*
logarea	0,1511	-16,77%	0,1343	-26,02%	-11,11%
proprietario			0,2195		
no obser	18547		18547		
R2	0,4110		0,4156		
R2 ajustado	0,4104		0,4149		

FONTE: Elaboração dos autores.
Itálicos indicam diferenças expressivas em relação ao estudo de comparação.

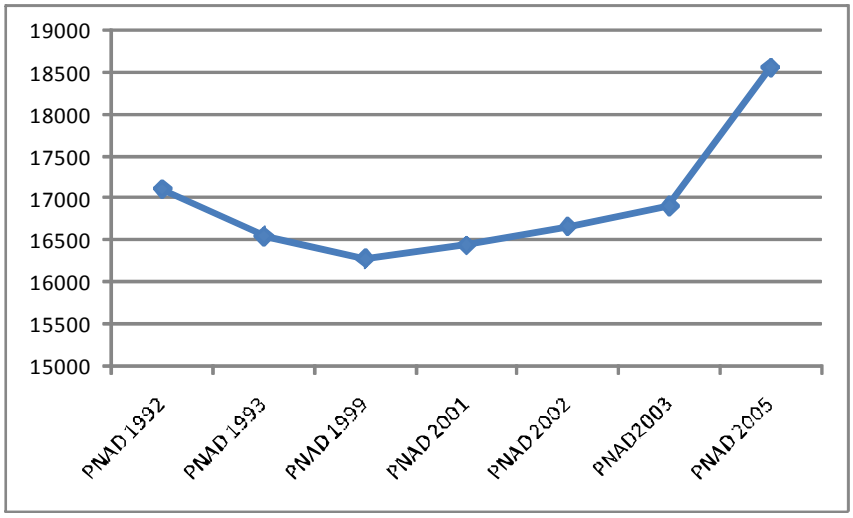
Tínhamos visto que as maiores diferenças entre os coeficientes estimados nas regressões, entre nossas estimativas da PNAD 2001 e as de Hoffmann e Ney (2003), correm nas variáveis contapp (conta própria), empredor (empregador) e indigena (indígena). Esses valores mais discrepantes estão representados em itálico. Para os outros anos, em 1992 as mesma variáveis apresentam coeficientes bem diferentes em relação ao estudo apontado. E também o intercepto: o que era de se esperar por causa da inflação entre os períodos. Na PNAD de 1993, nota-se grande diferença nos coeficientes, em relação ao estudo de comparação, para as mesmas variáveis do ano anterior, além de idade, idade2 e amarela (números em itálico). Em 1999, 2001, 2002 e 2005, só se destacam como bem diferentes os coeficientes das mesmas três variáveis de 1992; agora, em todos esses casos, o coeficiente do intercepto é próximo ao caso comparado. Então percebe-se que os casos mais discrepantes em relação a Hoffmann e Ney (2003), coeficientes das binárias conta própria, empregador e indígena, se mantém em todas as PNADs analisadas, para todos esses anos.

O efeito da inclusão da variável "proprietário"

Ainda analisando os resultados das regressões mostrados na tabela 4, vejamos como se modificam os coeficientes das variáveis nas equações de rendimento quando se inclui a variável que capta a condição de proprietário da terra rural. Primeiramente, cabe destacar que o R2 ajustado aumenta em todas as PNADs consideradas com a inclusão da variável "proprietário". Isso mostra que a nova variável melhora a qualidade da

estimação.[7] Outro aspecto que queremos destacar é que o tamanho da chamada PNAD rural (pessoas ocupadas na agricultura como atividade principal) altera-se pouco nos anos considerados (Gráfico 1). Observa-se que o tamanho da amostra da PNAD rural começa com pouco mais de 17 mil observações em 1992, cai para pouco menos de 16.500 pessoas em 1999 e na década seguinte é crescente até cerca de 18.500 observações em 2005.

Gráfico 1 – Tamanho da amostra para a PNAD rural – número de observações



A inclusão da variável de propriedade mostrou-se então desejada em todos os casos. Vejamos como tal fato altera o efeito de cada uma das variáveis consideradas sobre o logaritmo do rendimento. Em algumas variáveis, a inclusão da *proxy* de propriedade torna mais intenso o efeito delas, casos de idade, idade 2, sexo, amarela, loghtrab, regiões do país e domicílio para as PNAD de 1992; sexo, logtrab, todas as regiões, exceto sul, e domicílio para a PNAD de 1993, idade2, sexo, indigena, loghtrab, SP, CO e domicílio em 1999; idade2, sexo, amarelo, loghtrab, condunidom, SP e domicílio em 2001; idade, idade2, sexo, loghtrab, CO e domicilio em 2002; sexo, loghtrab, condunidom, SP e domicilio em 2003; e idade 2, sexo, loghtrab e domicilio para a PNAD de 2005. Portanto, apenas sexo, loghtrab e domicílio sempre apresentam intensificados os efeitos delas sobre o logaritmo do rendimento quando se inclui a variável propriedade. Nos outros casos anteriormente omitidos, o efeito da propriedade é o de diminuir o impacto das variáveis na determinação da renda agrícola, medido pela queda do coeficiente de regressão.

É importante observar o comportamento dos coeficientes de escolaridade e de área (em logaritmo: logarea) quando se inclui a variável para propriedade. Em todos os casos analisados, o coeficiente de logarea se reduz significativamente com a inclusão da nova variável no modelo de regressão. A redução no coeficiente vai de 8 a 11% aproximadamente, dependendo do ano considerado. O coeficiente da regressão para a variável proprietário é sempre positivo: 0,2165 em 1992, 0,29 em 1993, 0,1881 em 1999, 0,1933 em 2001, 0,1792 em 2002, 0,1917 em 2003 e 0,2195 em 2005. Em todos os casos, ele é superior ao coeficiente do logaritmo da área.

A leitura dos coeficientes da regressão leva a concluir que a variável propriedade é um indicador melhor do impacto da riqueza do indivíduo na formação da renda

[7] Veja, a respeito, Maddala (p.124-127). As diferenças percentuais no R2 ajustado com a inclusão da variável "proprietario" estão representadas no gráfico 5, mais adiante.

agrícola do que a variável logaritmo da área. Além disto, nota-se que em 2001 as estimativas de Hoffmann e Ney (2003) dos coeficientes da área eram exageradas, isso à luz do nosso resultado para este ano, por não incorporar a variável proprietário. Mesmo nossas estimativas para essa variável de área, precisaram ser corrigidas e atenuadas com a introdução de uma *proxy* adicional.

O estudo de Hoffmann e Ney provavelmente estava exagerando a contribuição da área como medida de riqueza porque a área correlaciona-se com uma variável omitida que é a propriedade da terra. Nas áreas maiores, cresce a ocorrência de proprietários. Então o efeito que ele captou não era apenas um efeito de posse, mas também de propriedade da terra (esta última entendida como uma melhor medida de riqueza).

Comparando-se grupos de proprietários e de não proprietários

Um exercício que se mostrou esclarecedor consiste em separar a PNAD rural em dois grupos de proprietários e de não proprietários. Com o uso do SAS como ferramenta, fica fácil desmembrar a base de dados nesses dois subconjuntos. Tendo-se criado anteriormente a variável "proprietario" com base na variável original da PNAD V9021 (condição do entrevistado em relação ao empreendimento agrícola), o que temos que fazer é pedir para o programa apagar todas as observações nas quais a referida variável assume o valor 0, isto é, todos os casos em que a pessoa é não proprietária. Evidentemente o outro conjunto, o de não proprietários, é gerado simplesmente com o comando que suprime as observações em que a mesma *proxy* assume o valor 1. Na recém criada base de dados exclusiva para proprietários, como os proprietários são apenas de dois tipos, os empregadores e os de conta própria (não podem ser empregados), a informação da variável "contapp" já gera a informação do conteúdo de empredor: o valor do intercepto e o da variável "contapp" geram o valor assumido por empredor, isto é, uma variável é combinação linear de duas outras, o que torna a inclusão dela inapropriada em um modelo de regressão. O SAS simplesmente assume valor zero para a variável linearmente dependente, ou então eliminamos a variável "empredor" (ou "contapp") na estimativa da equação de rendimento. Note que agora as variáveis explicativas não incluem "empredor" (e nem "proprietario", pois todas as observações são para proprietários).

A tabela 5 mostra os coeficientes das variáveis em equações de rendimento estimadas para toda as PNADs que estamos considerando (1992, 1993, 1999, 2001, 2002, 2003 e 2005). Na primeira coluna aparecem as mesmas variáveis avaliadas na PNAD 2001. Na segunda coluna, são estimados os coeficientes das variáveis (e do intercepto) na estimativa para a base de proprietários. A próxima coluna evidencia como os coeficientes se alteram em relação ao modelo básico aplicado a toda a PNAD rural (neste caso, também com a variável "empredor"). Conforme se vê, alguns coeficientes crescem e outros decrescem, em relação ao modelo básico, quando se considera apenas a PNAD rural de proprietários. Entre outras informações, nota-se que o coeficiente para escolaridade, na nova situação, cai bastante nas PNADs de 1992 (-24,5%), 1993 (-22%). Ele aproximadamente se mantém para 1999 (+2,8%), 2001 (-5,2%), 2002 (+6%); aumenta muito em 2003 (+16,8%) e significativamente em 2005 (+8,45%). Ou seja, não se observa uma relação estável ao longo do tempo entre ser proprietário e ter um efeito maior da variável escolaridade na determinação da renda.

Observa-se também os coeficientes para o logaritmo da área (variável "logarea"). A área mostra-se menos importante na determinação da renda em 1993 (-13%), 1999 (-11%); mais importante em 2001 (+11,2%) e quase não se altera o seu

coeficiente se o grupo é o de proprietários em 1992, 2002, 2003 e 2005. Então é importante a leitura das variações no coeficientes nessa comparação entre o modelo de regressão para a PNAD rural como um todo e o caso somente com proprietários. Dificilmente os coeficientes mudariam de modo sistemático. O tamanho da área, ou a escolaridade ou outras das variáveis consideradas estão razoavelmente correlacionados como o fato da pessoa ser proprietária: os proprietários em geral ocupam áreas maiores e tem um grau maior de escolaridade. Mas a mudança no nível das variáveis não implica em coeficientes maiores na equação de regressão, pois a interpretação do coeficiente é mais no sentido da contribuição marginal de acrescimentos nos valores das variável do que no efeito dos montantes absolutos. Exemplificando, quando se alcança certo nível de escolaridade, um ano a mais de estudo passa a fazer menos diferença na formação da renda do que um ano a mais de escolaridade para um semi-analfabeto. A combinação de efeitos de escala e de efeitos nas variações marginais faz com que os coeficientes das variáveis ora aumentem, ora decaiam, dependendo do ano da PNAD considerada, quando se compara os proprietários com o grupo misto.

Note também a quarta coluna da tabela 5 com os coeficientes para a PNAD rural de não proprietários. Cabe comparar estes resultados com o modelo anterior (proprietários). O intercepto sempre cai em todos os anos analisados, mostrando que os não proprietários são mais pobres. O coeficiente de área é bem menor para os não proprietários (queda entre 31 e 59%, dependendo do ano), o que corrobora a hipótese de Hoffmann e Ney (2003) de que a área esteja correlacionada com a posse da terra. Já a diferença no efeito da escolaridade na determinação da renda flutua muito, dependendo da PNAD em questão, sendo ora maior no grupo de não proprietários (1992 e 1993), ora consideravelmente menor nesse caso, como nos demais anos considerados. Na maioria das PNADs observadas, o efeito da escolaridade é bem menor no caso dos não proprietários. Tal constatação pede uma reflexão teórica: por que anos a mais de estudo tendem a ter menor impacto na formação da renda agrícola no caso de não proprietários? Uma resposta que só podemos especular. Talvez porque o nível de escolaridade seja muito baixo nas pessoa deste grupo de não proprietários e, nesse caso, importa mais na geração da renda a saúde e a disposição para o trabalho do que a educação formal, que poderia acrescentar pouco na capacidade produtiva nesses baixos níveis de escolaridade.

Tabela 5 – Equações de rendimento estimadas para pessoas ocupadas na agricultura brasileira. Coeficientes de regressão para a PNAD rural de proprietários e de não proprietário. PNAD: 1992, 1993, 1999, 2001, 2002, 2003 e 2005[1]

Variável	Modelo 1 - PNAD rural de proprietários		Modelo 2 - PNAD rural de não proprietários	
	Coeficientes	Variação em relação ao modelo para toda a PNAD rural	Coeficientes	Variação em relação ao Modelo 1
PNAD 1992				
intercepto	10,4931	2,61%	9,5119	-9,35%
escolaridade	0,0526	-24,54%	0,0748	42,30%
idade	0,2810	-22,22%	0,4275	52,14%
idade2	-0,0311	-19,60%	-0,0489	56,97%
sexo	-0,1824	-20,52%	-0,2499	36,96%

contapp	-0,7018	-47,62%	-0,8225	17,19%
preta	-0,0413	-51,40%	-0,0765	85,19%
indígena	-0,1489	-62,59%	-0,4006	*169,10%*
parda	-0,0582	-25,45%	-0,0748	28,43%
amarela	0,4519	2,68%	0,4103	-9,21%
loghtrab	0,2749	-35,00%	0,5805	*111,19%*
condunidom	0,2019	3,42%	0,1810	-10,35%
NO	-0,0964	*-158,86%*	0,2400	*-348,96%*
SE	0,2509	13,06%	0,2157	-14,05%
SP	0,6834	15,70%	0,5725	-16,22%
SUL	0,4016	16,97%	0,3259	-18,84%
CO	0,2890	-15,69%	0,3591	24,25%
domicilio	-0,2757	*143,75%*	-0,0810	-70,61%
logarea	0,1353	*4,64%*	0,0699	*-48,34%*
No. observações	4923		12187	
R2	0,3740		0,3386	
PNAD 1993				
intercepto	6,4046	-2,82%	6,3306	-1,16%
escolaridade	0,0530	-21,96%	0,0700	32,16%
idade	0,2735	*1268,62%*	0,0256	*-90,66%*
idade2	-0,0309	*11037,16%*	-0,0003	*-98,97%*
sexo	-0,3840	74,65%	-0,2086	-45,69%
contapp	-0,7067	-52,07%	-0,9890	39,96%
preta	-0,0671	-30,59%	-0,0846	26,13%
indígena	-0,6997	90,46%	-0,2093	-70,09%
parda	-0,1155	3,55%	-0,0953	-17,46%
amarela	0,1478	-9,41%	0,1917	29,68%
loghtrab	0,3617	-22,81%	0,5205	43,91%
condunidom	0,2049	-44,25%	0,3602	75,77%
NO	0,2014	-49,78%	0,4839	*140,24%*
SE	0,1666	-25,95%	0,2568	54,14%
SP	0,6954	2,95%	0,6954	-0,01%
SUL	0,6215	16,42%	0,4914	-20,93%
CO	0,4636	-13,31%	0,5807	25,27%
domicilio	-0,2072	*158,44%*	-0,0569	-72,54%
logarea	0,1340	*-12,92%*	0,0924	*-31,04%*
No. observações	4800		11754	
R2	0,3936		0,3657	
PNAD 1999				
intercepto	3,0839	12,18%	2,1201	-31,25%
escolaridade	0,0620	2,75%	0,0542	-12,70%
idade	0,2803	-18,97%	0,4002	42,77%
idade2	-0,0324	-16,82%	-0,0475	46,67%
sexo	-0,2507	8,42%	-0,2359	-5,91%
contapp	-0,7586	-50,76%	-1,0447	37,72%
preta	-0,1343	9,02%	-0,1050	-21,84%

indígena	-0,3859	22,56%	-0,3106	-19,52%
parda	-0,1803	38,02%	-0,0986	-45,29%
amarelo	0,2140	-37,01%	0,3565	66,57%
loghtrab	0,1961	-43,57%	0,4848	*147,28%*
condunidom	0,1654	-9,21%	0,1651	-0,18%
NO	0,0541	-68,95%	0,2237	*313,80%*
SE	0,1616	-21,74%	0,2352	45,50%
SP	0,7095	*27,96%*	0,5572	-21,46%
SUL	0,2071	1,81%	0,2154	3,99%
CO	0,2720	-20,68%	0,3829	40,76%
domicilio	-0,2482	*135,13%*	-0,0630	-74,60%
logarea	0,1266	*-10,98%*	0,0864	*-31,79%*
No. observações	5133		11155	
R2	0,4234		0,3889	
PNAD 2001				
intercepto	2,5969	-0,79%	1,9776	-23,85%
escolaridade	0,0593	-5,20%	0,0591	-0,32%
idade	0,2285	-36,72%	0,4012	75,61%
idade2	-0,0280	-27,70%	-0,0443	58,30%
sexo	-0,2478	8,23%	-0,2338	-5,67%
contapp	-0,6628	-63,49%	-1,0148	53,11%
preta	-0,1317	*31,21%*	-0,0825	-37,35%
indígena	-0,7503	482,83%	-0,0880	-88,27%
Amarelo	0,5595	-9,03%	0,5687	1,65%
parda	-0,1161	7,98%	-0,0926	-20,30%
loghtrab	0,1907	-48,97%	0,5136	*169,33%*
condunidom	0,1701	-12,02%	0,1806	6,23%
NO	-0,0695	*-129,67%*	0,3574	*-614,44%*
SE	0,1557	-26,03%	0,2296	47,47%
SP	0,7139	27,13%	0,5609	-21,43%
SUL	0,3810	9,81%	0,3197	-16,09%
CO	0,3209	-33,37%	0,5329	66,05%
domicilio	-0,2887	*199,99%*	-0,0348	-87,95%
logarea	0,1771	*11,16%*	0,0731	*-58,76%*
No. observações	5124		11322	
R2	0,3922		0,3745	
PNAD 2002				
intercepto	3,0384	18,68%	1,7932	-40,98%
escolaridade	0,0628	5,94%	0,0533	-15,09%
idade	0,2634	-22,27%	0,3852	46,24%
idade2	-0,0306	-16,34%	-0,0437	42,86%
sexo	-0,2583	38,58%	-0,1855	-28,18%
contapp	-0,9199	-47,85%	-1,2218	32,81%
preta	-0,1245	*117,98%*	-0,0155	-87,56%
indígena	-0,4307	66,76%	-0,1159	-73,09%
Amarelo	0,1505	-50,94%	-0,5344	-455,17%

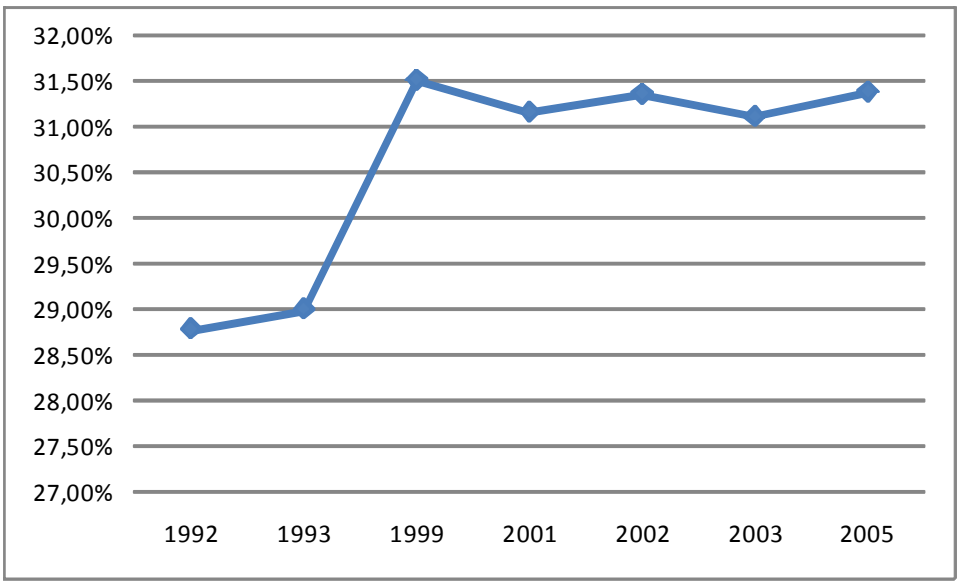
parda	-0,1245	59,28%	-0,0472	-62,07%
loghtrab	0,1890	-55,23%	0,5982	*216,53%*
condunidom	0,1622	-23,43%	0,2011	23,95%
NO	-0,0311	*-118,68%*	0,2565	*-925,12%*
SE	0,2551	6,44%	0,2295	-10,05%
SP	0,6689	24,09%	0,5423	-18,92%
SUL	0,4916	12,24%	0,4007	-18,49%
CO	0,2943	-34,30%	0,5074	72,41%
domicilio	-0,2377	*145,97%*	-0,0478	-79,90%
logarea	0,1551	*0,12%*	0,0950	*-38,74%*
No. observações	5228		11443	
R2	0,4362		0,3991	
		PNAD 2003		
Intercepto	2,4701	-13,60%	2,2643	-8,33%
escolaridade	0,0638	16,81%	0,0267	-58,18%
idade	0,3419	*1159,12%*	0,0267	-92,20%
idade2	-0,0349	*13610,03%*	-0,0003	-99,28%
sexo	-0,3037	47,54%	-0,2008	-33,89%
contapp	-0,7536	-55,72%	-1,0599	40,65%
preta	-0,2426	69,62%	-0,0958	-60,50%
indigena	-0,3431	5,03%	-0,2540	-25,97%
parda	-0,1288	25,00%	-0,0750	-41,78%
Amarela	0,3266	-2,30%	0,1088	-66,70%
loghtrab	0,2596	-50,12%	0,6761	*160,41%*
condunidom	0,1838	-26,04%	0,2283	24,25%
NO	0,0604	-78,42%	0,3665	*506,80%*
SE	0,2607	-5,49%	0,2779	6,61%
SP	0,9180	48,78%	0,5914	-35,58%
SUL	0,5873	14,87%	0,4183	-28,76%
CO	0,3309	-33,61%	0,5642	70,48%
domicilio	-0,1252	*162,23%*	-0,0199	-84,08%
logarea	0,1444	*-2,62%*	0,0781	*-45,92%*
no obser	5261		11649	
R2	0,4106		0,3917	
		PNAD 2005		
intercepto	2,7401	6,43%	2,0285	-25,97%
escolaridade	0,0638	8,45%	0,0513	-19,61%
idade	0,3419	-2,48%	0,4263	24,69%
idade2	-0,0349	-0,77%	-0,0467	33,74%
sexo	-0,3037	50,19%	-0,2089	-31,20%
contapp	-0,7536	-59,01%	-1,0962	45,47%
preta	-0,2426	656,11%	-0,0423	-82,56%
indigena	-0,3431	29,08%	-0,2939	-14,34%
parda	-0,1288	155,22%	-0,0449	-65,13%
Amarela	0,3266	161,31%	0,2038	-37,59%
loghtrab	0,2596	-44,86%	0,5965	*129,73%*

condunidom	0,1838	-3,57%	0,1506	-18,02%
NO	0,0604	-85,66%	0,4282	*608,86%*
SE	0,2607	-12,09%	0,2783	6,74%
SP	0,5873	-10,67%	0,6551	11,55%
SUL	0,5873	14,67%	0,4471	-23,87%
CO	0,3309	-39,62%	0,6230	88,26%
domicilio	-0,1252	*548,91%*	0,0047	*-103,76%*
logarea	0,1444	*-4,41%*	0,0704	*-51,23%*
no obser	5821		12726	
R2	0,4106		0,4144	

[1] Itálico identifica alterações importantes
FONTE: Elaboração do autor

Em alguns poucos casos, os coeficientes não deram significativos (em nível de 5% de significância), como em "preto" e "norte" para 2001, contudo, a tabela anterior não indica essas ocorrências. O parâmetro R2 da regressão piora no caso de não proprietários (apenas em 2005 ele se mantém), o que isoladamente considerado poderia indicar a qualidade maior do modelo para o caso de proprietários e sua maior inadequação para a base de não proprietários, muito embora esse parâmetro isoladamente não permita conclusões cabais a esse respeito. Por fim, gostaríamos de chamar a atenção para a distribuição da amostra entre proprietários e não proprietários. O número de observações no grupo de não proprietários oscila entre 11 mil e quase 13 mil pessoas; o grupo dos proprietários gira em torno de cinco mil observações (quase seis mil em 2005). Portanto, há bem mais não proprietários do que proprietários. O gráfico 2 mostra que os proprietários representa cerca de 30% da amostra. O número de proprietário cresceu significativamente entre as PNADs de 1993 e 1999 e depois se estabilizou no novo patamar.

Gráfico 2 - Porcentagem de proprietários na amostra da PNAD rural: 1992, 1993, 1999, 2001, 2002, 2003 e 2005

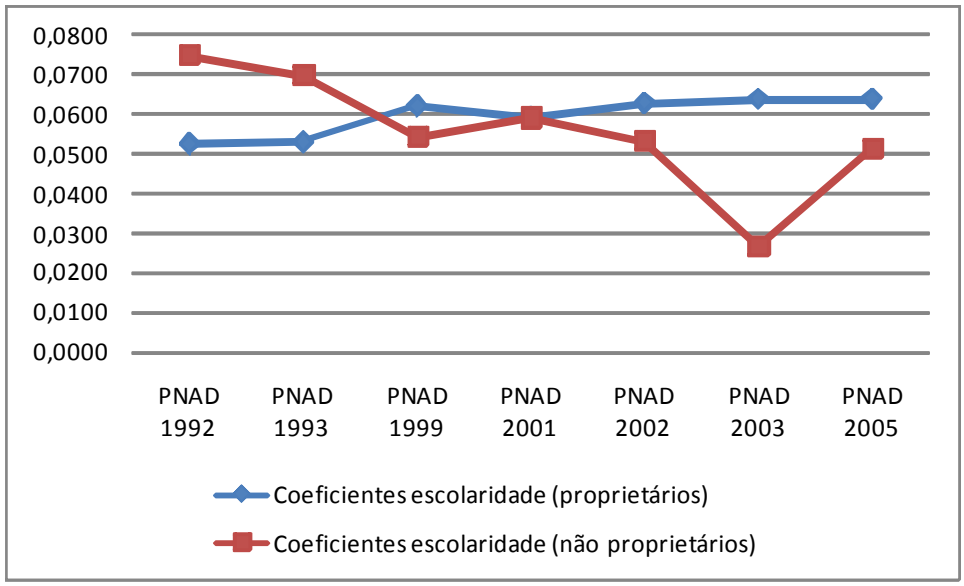


O crescimento do número relativo de proprietário na amostra, entre as PNADs de 1993 e 1999, muito provavelmente deve estar relacionado com o início de uma onda

importante de reforma agrária no país que se iniciou a partir de 1995. Com os assentamentos das famílias de sem-terras, um número maior de produtores passaram da condição de não proprietários (em geral empregados ou conta próprias) para a condição de proprietários, dada a titularidade da terra. Isso refletiu na composição da PNAD rural, embora o efeito seja tímido, pois a participação dos grupos não flutuou mais do que 2% entre 1992 e 2005.

Cabe retomar a discussão anterior sobre o comportamento das variáveis escolaridade e área (logaritmo da área) na determinação da renda agrícola quando se compara os dois grupos que estamos a considerar. O gráfico 3 permite uma visualização completa do que ocorreu com o coeficiente de escolaridade nesses grupos. Conforme já antecipamos, o coeficiente começa acima no grupo de não proprietários, porém, já a partir de 1999 ele perde para o coeficiente da mesma variável no modelo aplicado a proprietários. Nota-se que a tendência para o grupo de não proprietários foi a de queda quase continua do coeficiente até 2003, com aumento em 2005. Que dizer: a variável escolaridade torna-se, por alguma razão, bem menos impactante na estimativa da renda nas últimas PNADs analisadas, no caso dos não proprietários. Já no grupo dos proprietários, o coeficiente estimado para escolaridade é bem mais estável, com ligeira tendência de aumento ao longo das PNADs que se sucedem no tempo. É mesmo de se esperar que os coeficientes sejam mais expressivos no grupo dos proprietários: por serem mais ricos (compare os interceptos), os proprietários, em média, já atingiram um patamar a partir do qual a educação passa a importar mais na geração da renda, ao contrário do outro grupo em que muito de seus membros estão numa condição na qual déficit de saúde etc. são mais críticos do que a falta de instrução.[8]

Gráfico 3 – Evolução dos coeficientes da variável "escolaridade". Grupo de pessoas ocupadas na agricultura proprietárias e não proprietárias. PNAD rural: 1992, 1993, 1999, 2001, 2002, 2003 e 2005

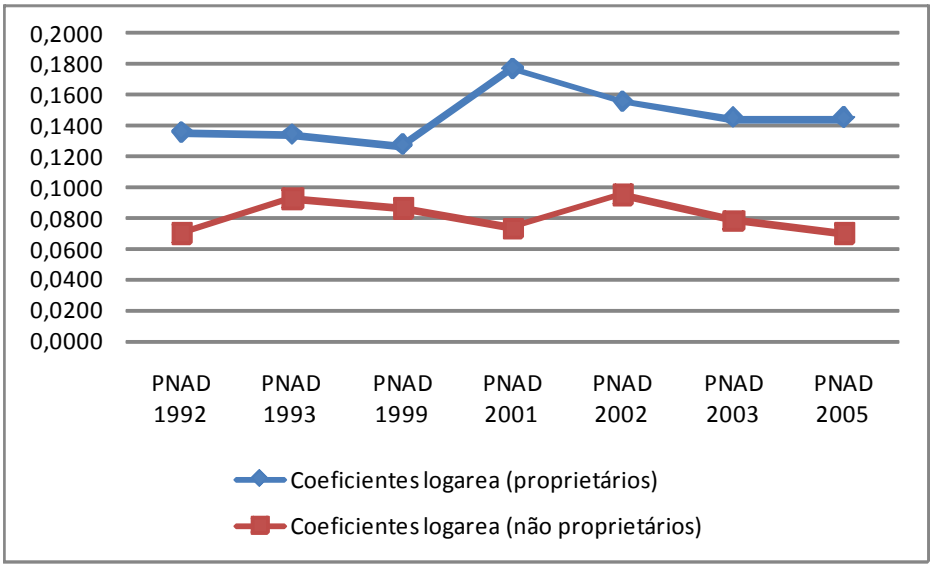


O gráfico 4 mostra a evolução dos coeficientes para a variável "logarea" para as PNADs consideradas. Chama a atenção o fato de o coeficiente estimado para o logaritmo da área ser sistematicamente superior no grupo dos proprietários. Este resultado ressalta diretamente um aspecto do trabalho de Hoffmann e Ney (2003): a

[8] Não é ocioso lembrar, mais uma vez, de que as comparações entre resultados de PNADs de diferentes anos devem ser vistas com muita precaução em face da ligeira mudança metodológica entre os anos e o fato de as amostras não serem exatamente as mesmas.

estimativa para o coeficiente de área que eles obtiveram superestimou a contribuição desta variável na determinação da renda agrícola, pois, consideraram a PNAD rural com proprietários e não proprietários. Se tivessem considerado apenas a PNAD rural para não proprietários, o valor do coeficiente dessa variável teria sido consideravelmente menor. A diferença entre as trajetórias deste coeficiente entre proprietários (sempre maior) e não proprietários é particularmente maior na PNAD analisada pelos autores, que foi a de 2001. Note, no gráfico 4, que a distância vertical entre as duas curvas é máxima no ano de 2001. A explicação do viés evidente, e não comentado pelos autores, é a de que a variável área (ou o log dela) está correlacionada com a variável omitida "proprietário", de forma que o coeficiente da primeira fica inflado pelo efeito embutido da condição de ser proprietário, que assim como a área contribui para ganhos de rendimentos agrícolas: a segunda pelos efeitos econômicos dos direitos de propriedade e a primeira, como apontaram os autores, como efeito da posse da terra: medida de riqueza segundo eles.

Gráfico 4 – Evolução dos coeficientes da variável "logarea" (logaritmo da área do estabelecimento). Grupo de pessoas ocupadas na agricultura proprietárias e não proprietárias. PNAD rural: 1992, 1993, 1999, 2001, 2002, 2003 e 2005



Em suma, parece-nos evidente a vantagem de separar a PNAD rural em dois grupos, definidos pelo fato de a pessoa ser proprietária ou não, e replicar o modelo de regressão de Hoffmann e Ney (2003). Também nos parece revelador comparar os resultados da regressão nos grupos separados, um com o outro, e também o resultado obtido em cada caso com as estimativas feitas naquele estudo. Embora as contribuição dos autores tenha sido da maior relevância, pensamos que o avanço da pesquisa sobre o efeito da posse da terra demanda mesmo uma análise comparativa ente PNADs de diferentes anos, e ainda, a incorporação explícita da variável de propriedade separando, destarte, efeitos de posse da terra, de um lado, e da propriedade dela, de outro, na determinação da renda agrícola.

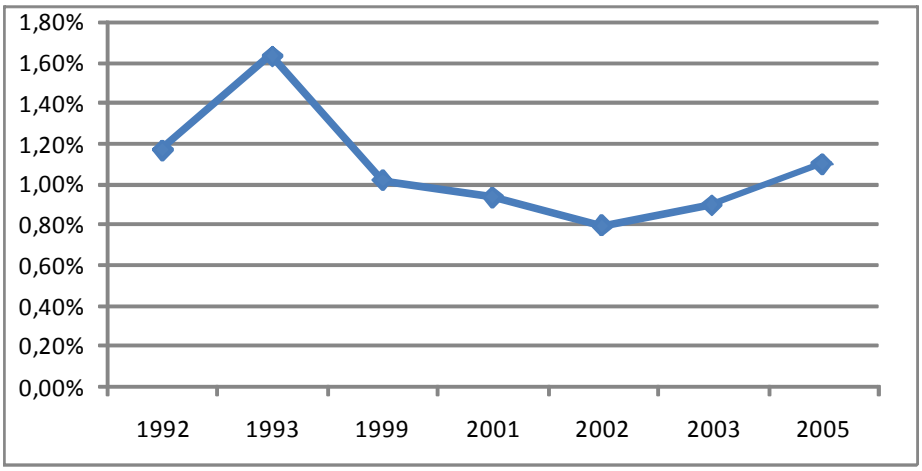
Interpretação dos resultados e comparação com as expectativas geradas com base em considerações teóricas

Vejamos nesta seção o que os resultados observados com o tratamento de dados das PNADs nos dizem à luz das expectativas criadas pela exposição teórica em seção anterior. Com dissemos na seção teórica, três conclusões eram esperadas da parte empírica do trabalho:

1. Esperávamos que a incorporação da variável "proprietário" melhorasse a regressão (medindo-se a qualidade dela e a intensidade da melhora pelo crescimento do R2 ajustado).
2. Que, sob o impacto da reforma agrária acentuada desde 1995, as diferenças nas médias de rendimento entre os grupos de proprietário e de não proprietários estivessem diminuindo.
3. Que os coeficientes de regressão da variável "proprietário" (para toda a PNAD rural) crescessem ao longo dos anos, sendo maior nas PNADs mais recentes (indicativo das novas possibilidades produtivas trazidas pela titularidade da terra).

O gráfico 5 mostra que o primeiro resultado esperado, de fato, foi alcançado: a incorporação da variável melhorou a regressão medida pelo crescimento do R2 ajustado[9]. Note que as diferenças percentuais nesse parâmetro de regressão são sempre maiores que 0,8% em relação ao modelo sem a variável proprietário. Em 1993, o diferencial chega a mais de 1,6%, e a menor distância ocorre em 2002. Em 2001, ano em que recai os estudos de Hoffmann e Ney (2001), a incorporação da variável proprietário melhorou o R2 ajustado em quase 1%. Esses resultados evidenciam a relevância de se considerar uma *proxy* binária específica indicadora de propriedade: melhoram as estimativas da renda esperada em relação ao modelo que considera apenas indicadores de posse da terra tais como ocupação e área do estabelecimento.

Gráfico 5 – Evolução do parâmetro de regressão R2 ajustado. Diferenças percentuais nesse parâmetro com incorporação da variável "proprietário". PNAD rural: 1992, 1993, 1999, 2001, 2002, 2003 e 2005



A segunda conclusão esperada, a de que as diferenças nas média de rendimento entre os grupos de proprietários e de não proprietários estivessem diminuindo, tal resultado também foi alcançado observando-se as informações oriundas das PNADs. As diferenças percentuais, para a PNAD de 1992, entre os rendimentos estimados de

[9] O crescimento no R2 ajustado não é um critério exaustivo para se avaliar melhorias na qualidade da regressão. Também sabemos que a observação exclusiva deste indicador pode levar a conclusões equivocadas.

proprietário e de não proprietários (a diferença entre rendimentos da categoria de proprietários e da categoria de não proprietários, tomada como base) foi de 65,97%. Os mesmo resultado para a PNAD de 2005 foi de 42,73%. Portanto, a renda esperada para proprietários é significativamente superior à dos não proprietários, e as diferenças, conforme esperado, diminuíram entre 1992 (período anterior à forte expansão da reforma agrária) e o ano de 2001 (já sob o impacto dos esforços cumulativos de assentamentos de milhares de famílias rurais).

Foi feito o cálculo da renda esperada, isto é, as diferenças percentuais, para a variável binária considerada no modelo de regressão, entre os rendimentos estimados, incluindo-se a binária em questão, e os rendimentos estimados da categoria tomada como base (todos os anos em questão). Vejamos, em pormenores, o método utilizado no cálculo das diferenças na renda esperada. A equação 1 mostra o modelo de regressão utilizado para a estimativa da equação de rendimento para pessoas ocupadas na agricultura brasileira:

$$(1) \quad Y_j = \alpha + \Sigma_i \beta_i X_{ij} + \beta_E E_j + \beta_j I_j + \beta_{j2} I_j^2 + \beta_T \ln T_j + \beta_W \ln(A_j) + u_j$$

X_{ij}'s representam todas as variáveis binárias consideradas na equação de regressão (inclusive "proprietário"), E_j é a escolaridade, I_j representa a idade da pessoa[10], T_j é o número de horas semanais de trabalho e A_j capta a área do estabelecimento rural. Os β's representam os coeficientes a serem estimados e α é o intercepto. Aplicando-se o operador esperança E nos dois membros da equação chega-se ao valor da renda esperada $E(Y_j)$ (equação 2):

$$(2) \quad E(Y_j) = E(\alpha) + \Sigma_i E(\beta_i X_{ij}) + E(\beta_E E_j) + E(\beta_j I_j) + E(\beta_{j2} I_j^2) + E(\beta_T \ln T_j) + \\ + E(\beta_W \ln A_j) + E(u_j)$$

Colocando-se os coeficientes fora do operador esperança (por serem valores fixos), e levando-se em conta o coeficiente α fixo e erros aleatórios heterocedásticos com as propriedades usuais ($E(u_j) = 0$), chega-se à equação 3:

$$(3) \quad E(Y_j) = \alpha + \Sigma_i \beta_i E(X_{ij}) + \beta_E E(E_j) + \beta_j E(I_j) + \beta_{j2} E(I_j^2) + \beta_T E(\ln T_j) + \\ + \beta_W E(\ln A_j)$$

Com base na equação 3, estimamos os rendimentos esperados da categoria considerada, isto é, considera-se a PNAD rural que só contenha exemplares da categoria em questão (todos negros ou todos sendo mulheres etc.), e compara-se com os rendimentos estimados da categoria tomada como base (PNAD rural só com brancos ou só com homens etc.). Para cada *proxy* binária considerada, não alteramos as *proxies* para outras variáveis, porém, no grupo da categoria em questão eliminamos do modelo de regressão as *proxies* relacionadas a esse grupo. Por exemplo, na estimativa para a renda esperada em diferentes cores, eliminamos as *proxies* de cor e rodamos o programa que gera a regressão aplicado ora na base de dados da categoria em questão (cuja diferença de rendimentos em relação à categoria base queremos estimar), ora na categoria de base e verificamos a diferença percentual nas duas estimativas da renda.

Portanto, para cada variável binária gera-se a sub base da PNAD rural desta categoria, e monta-se a regressão para a estimativa dos coeficientes. Depois, calcula-se os valores médios das variáveis do modelo (média para a sub base em questão) e

[10] Respeitadas as convenções que tem sido adotadas: considerar idades em dezenas de anos e escolaridade como sendo anos de estudo até 14 anos de escolaridade, e mais do que isso assume 17.

multiplica-se o coeficiente estimado pela respectiva média da variável, conforme indica a equação 3. Finalmente, soma-se todos os resultados do produto da variável por seu coeficiente, mais o valor do intercepto, e chega-se à estimativa do logaritmo da renda esperada para certa categoria. Calculando o número natural elevado ao valor do log da área, anula-se o log e obtém-se a renda esperada. Repete-se esse mesmo procedimento para todas as categorias em questão e compara-se todos os resultado para cada grupo de binárias com o correspondente valor da categoria tomada como base (branco, se variáveis de cor; homem, se variável sexo; não arrimo, se variável da condição na unidade familiar; Nordeste, se região; empregado, se condição na ocupação; não rural (ou urbano), se domicílio, e não proprietário, se for a variável "proprietário").

Então, trata-se de um procedimento demorado e trabalhoso. Fizemos apenas para duas PNADs: 1992, o começo da série, e 2001, mesma PNAD considerada por Hoffmann e Ney (2003). A tabela 6 mostra que, para a categoria cor, em relação aos brancos, os pretos são 42% mais pobres, os pardos 41% e os indígenas 46%, na PNAD de 1992. Já os amarelos (orientais) são, em média, 239% mais ricos que os brancos. Para sexo, em relação aos homens as mulheres são 41% mais pobres em rendimentos. Os arrimos de família são 64% mais ricos que os não arrimos; em relação ao nordeste, os agricultores do norte são 43% mais ricos, os do sudeste 56%, os de São Paulo 156%, os do sul 114% e os rurais do centro-oeste ganham 108% a mais que os nordestinos. Os de conta própria ganham 19% a mais do que os empregados e os empregadores ganham 118% acima do mesmo grupo base. Os rurais ganham 26% a menos que os urbanos e os proprietário ganham 66% a mais do que os não proprietários.

Tabela 6 – Comparações: renda esperada em relação à categoria base.

	Hoffmann e Ney (2003) PNAD 2001	PNAD 1992	PNAD 2001	Diferença entre nosso cálculo e o de Hoffmann e Ney (2003)	Diferenças entre a renda esperada das estimativas de 2001 em relação às de 1992
sexo	-14,9	-41,01	-36,39	-144,23%	11,27%
preta	-7,9	-41,75	-35,8	-353,16%	14,25%
parda	-9,8	-40,76	-35,92	-266,53%	11,87%
indígena	6,5	-46,53	-40,02	-545,48%	13,99%
amarela	74,2	239,01	171,41	131,01%	-28,28%
contapp	-30,8	18,55	-5,07	-83,54%	-72,67%
empredor	37,4	118,24	250,1	568,72%	111,52%
condunidom	22,1	64,21	44,88	103,08%	-30,10%
NO	27,6	43,04	77,27	179,96%	79,53%
SE	21,4	56,91	61,46	187,20%	8,00%
SP	75	156,84	157,54	110,05%	0,45%
SUL	38	114,03	109,8	188,95%	-3,71%
CO	60,1	108,07	142,07	136,39%	31,46%
domicilio	-9,4	-26,49	-0,2	-97,87%	99,24%
propriedade		65,97	42,73		-35,23%

FONTE: Elaboração do autor.

Vejamos agora os resultados para a PNAD de 2001 (mesmo método de cálculo). A tabela 6 evidencia que, para a categoria cor, em relação aos brancos, os pretos e os pardos são 36% mais pobres e os indígenas 40%. Já os amarelos (orientais) são, em média, 171% mais ricos que os brancos. Para sexo, em relação aos homens as mulheres são 37% mais pobres em rendimentos. Os arrimos de família são 45% mais ricos que os não arrimos; em relação ao nordeste, os agricultores do norte são 77% mais ricos, os do sudeste 61%, os de São Paulo 157%, os do sul 110% e os rurais do centro oeste ganham 142% a mais que os nordestinos. Os de conta própria ganham 5% a menos do que os empregados e os empregadores ganham 250% acima do mesmo grupo base. Os rurais, em média, ganham praticamente o mesmo que os urbanos e os proprietário ganham 42% a mais do que os não proprietários.

Os resultados da renda esperada para a PNAD de 2001 em relação à PNAD de 1992 mostram que, em relação aos rendimentos agrícolas, vem ocorrendo um aumento nas disparidades regionais, já que os rendimentos em relação aos nordestinos cresceram (exceto na região sul do país); uma diminuição nos diferenciais de renda entre cor e entre gênero. Os empregadores aumentaram a vantagem sobre os empregados, mas os de conta própria pioraram e ficaram agora abaixo da renda média deste últimos. A diferença de renda esperada entre arrimos e não arrimos de família diminuiu; diminuiu a desvantagem dos residentes em domicílio rurais em relação aos de domicílio urbano e a diferença entre proprietários e não proprietários é menor, conforme já tínhamos assinalado.

A última coluna da tabela mostra as diferenças entre a renda esperada das estimativas de 1992 e 2001: 2001 em relação a 1992. As diferenças percentuais *não* estão em módulo, de forma que o sinal é importante: quando a renda esperada da categoria é menor do que a renda da categoria base em questão, a melhoria na desigualdade implica em um diferencial de renda esperada menos negativo, e as diferenças entre os indicadores nas duas PNADs são expressas em percentuais positivos (indicador de melhora, ou menos desigualdade): caso de sexo, etnias (exceto amarelo) e domicílio. Casos em que a vantagem em relação à categoria base se ampliou ocorrem em empregador e regiões (exceto sul), a vantagem piora, mas continua existindo em 2001 para a categoria propriedade, arrimos de famílias e cor amarela.

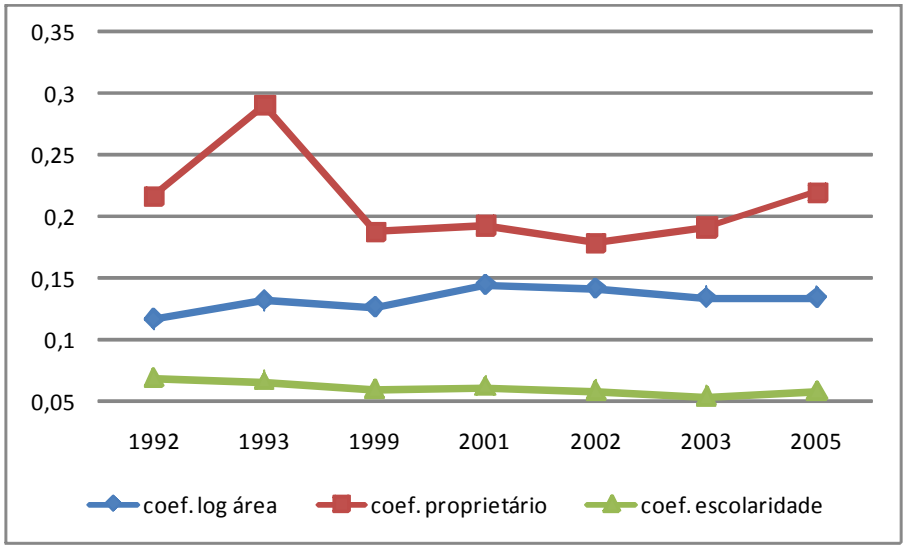
A tabela 6 (segunda coluna) também mostra as diferenças percentual entre os rendimentos estimados da categoria considerada e da categoria tomada como base que aparecem no artigo de Hoffmann e Ney (2003) para a PNAD de 2001. Os autores não especificam como a informação de renda esperada é gerada. Não sabemos se eles separam as PNADs rurais de acordo com as categorias; e se rodam regressões específicas com o fito de se fazer comparação, ou utilizam outro método. Com efeito, nota-se as diferenças obtidas pelos dois procedimentos de estimativas, o deles e o nosso. Por exemplo, enquanto para Hoffmann e Ney as mulheres ganham 15% a menos de renda estimada, para nós a defasagem de renda é ainda maior (36%). Os pretos para esses autores só ganham 8% a menos do que os brancos, enquanto em nosso cálculo a desvantagem é de 36%.

Pensamos que nossas estimativa dos diferenciais de renda esperada são bem mais plausíveis do que as de Hoffmann e Ney, por algum peculiaridade do método deles que desconhecemos. Esses autores parecem estar subestimando as diferenças das categorias em relação às respectivas categorias base. Por exemplo, não é muito provável que os rendimentos agrícolas em São Paulo sejam apenas 75% maior do que os do nordeste: o nosso número (157%) parece mais realista. Note que em todas as variáveis as disparidades em relação à categoria base aumentam em nossas estimativas em relação à referencia literária com a qual estamos comparando; exceto em domicílio e o curioso

caso da *proxy* indígena a qual nas estimativas dos autores teria uma renda agrícola, em média, maior do que as dos brancos, o que definitivamente não nos parece plausível.

Não se confirmou a expectativa de que os coeficientes de regressão da variável "proprietário" (para toda a PNAD rural) crescessem ao longo dos anos: indicativo das novas possibilidades produtivas trazidas pela titularidade da terra. De fato, o ano de 1999, já sob o efeito do surto de reforma agrária, apresenta um coeficiente de regressão para a variável em questão menor do que o de 1992 (gráfico 6). Daquele ano em diante, o referido coeficiente só cresce significativamente na PNAD de 2005. É o coeficiente da outra variável, o log da área, que a longo das PNADs consideradas cresce significativamente: do patamar de 0,1 para o patamar de 0,15. Esses resultados comparativos parecem traduzir o pouco efeito produtivo da reforma agrária: a condição de ser proprietário da terra não se traduz em um fator muito impactante no desempenho produtivo, na medida em que mesmo com o surto de assentamentos dos rurais e a concomitante titulação da terra a inclusão da variável propriedade ficou até menos impactante na equação de rendimentos do que era em 1992, antes da onda de assentamentos.

Gráfico 6 – Evolução dos coeficientes de logarea (logaritmo da área), proprietário e escolaridade nas equações de rendimento estimadas com inclusão da variável "proprietário". PNAD rural: 1992, 1993, 1999, 2001, 2002, 2003 e 2005



Talvez outros fatores estejam condicionando a determinação do coeficiente dessa variável, além da limitação metodológica nas comparações entre PNADs de diferentes anos. Contudo, os resultados obtidos nesta pesquisa parecem estar corroborando a tese de que a reforma agrária esteja sendo inócua do ponto de vista produtivo, em que pese a literatura em economia institucional argumentar a favor do impacto econômico dos direitos de propriedade.

A PNAD de 2001 apresentou o maior coeficiente para a variável "logarea". Ele foi menor tanto nas PNADs anteriores (1992, 1993 e 1999), quanto nas PNADs subseqüentes de 2002, 2003 e 2005. É difícil precisar o que tenha provocado oscilações nesses coeficientes, além das discrepâncias metodológicas e amostrais entre as PNADs. Se a variável área é entendida como medida de posse da terra e de riqueza, o efeito posse (riqueza no sentido de Hoffmann e Ney, mas não no nosso) parece ter sido mais forte justamente no ano em que os autores analisaram. Enfraquece-se um pouco suas

conclusões em outras PNADs, de outros anos. Por fim, note o ligeiro declínio do coeficiente de escolaridade nas equações de regressão para estimativa da renda agrícola. Ligeiramente, ao longo das PNADs analisadas, a variável escolaridade vem se tornando menos importante na estimação da renda esperada.

O uso do fator de expansão

Agora gostaríamos de comentar um detalhe técnico da pesquisa que já demandou investigação: o papel dos fatores de expansão da amostra e do peso das observações individuais na formação da amostragem. Os autores que tomamos como referência básica são explícitos em escrever que as equações de rendimentos das regressões utilizam, como ponderação para o método de mínimos quadrado, o fator de expansão associado a cada pessoa da amostra – Hoffmann e Ney (2003, p. 135). Provavelmente eles estão se referindo à variável PNAD V4729 que indica o chamado "peso da pessoa", tal como indicado no dicionário de variável da PNAD.

Há controvérsias quanto ao uso do fator de expansão da observação, ou do "peso da pessoa", nas estimadas de regressões que empregam dados da PNAD. Deparamo-nos com diversas opiniões técnicas a esse respeito, tanto do pessoal do IBGE quanto de colegas da FEA-RP/USP.[11] Alguns consideram essencial a inclusão do peso da observação individual na análise da regressão, enquanto outros nos alertaram quanto à distorção na análise trazida pela inclusão destes pesos. Há complexas questões econométricas envolvida que não pretendemos discorrer a respeito. Contudo, cabe um exercício ainda preliminar incorporando-se em nossas regressões o "peso da pessoa" com o intuito de se chegar a alguma conclusão inicial.

Discutem-se sobre a maneira de incorporar a variável V4729, da PNAD, no processo de análise estatística por regressão linear. O programa Stata possui um comando específico para a inclusão do referido peso. Não obstante, não conseguimos identificar na linguagem de programação do SAS nenhum comando ou rotina adequados para tratar da inclusão de pesos das observações específicas. Contudo, pretendemos fazer uma regressão pelo método ponderado usando a variável em questão. Para tanto, executamos um procedimento relativamente simples no qual foram geradas, na base da PNAD rural, colunas adicionais em que as variável explicativas, bem como a variável dependente do logaritmo da renda, são multiplicadas, em cada observação, pelo respectivo valor do "peso da pessoa" (variável V4729). No SAS, pode-se fazê-lo deixando o arquivo dos dados em modo de edição; a nova variável é criada simplesmente multiplicando-se a variável do modelo original de regressão V pelo pessoa da pessoa, resultando em nV. Em seguida, rodamos o programa com a nova rotina na qual, para efeito da regressão, multiplicam-se os dois membros da equação de estimação pelo "peso da pessoa". Este procedimento é usual na econometria. Fizemos alguns testes para avaliar a implicação deste método de ponderação.

Analisam-se então as diferenças nos coeficientes das variáveis e no teste de significância (teste t) em cada um dos modelos de regressão. Não nos interessa a mudança no intercepto da regressão, pois naturalmente ele foi multiplicado por um fator que traduz um peso médio das observações. A grande melhoria no R2, que foi de 0,39 no modelo original para 0,82 no caso do método ponderado, é sabidamente enganosa e não a levamos em conta. Importa-nos as mudanças nos coeficientes das variáveis. Também nos preocupamos com o teste t aplicado aos coeficientes das variáveis

[11] Agradecemos as sugestões dos professores Dr. Alexandre Nicolella e Dr. Luiz Guilherme Scorzafare a respeito desta questão.

explicativas no caso da regressão pelo método dos mínimos quadrados ponderados. Enquanto no modelo original apenas a estimação do coeficiente para a variável "indígena" não se mostrou significativa a um nível de 5%; no modelo ponderado pelos pesos, "preta", "parda" , "indigena" e região norte não tiveram coeficientes aceitáveis num nível de significância de 5%. Isto evidencia certa fraqueza das estimativas pelo método ponderado, que perde a significância para coeficientes em um número maior de casos.

A tabela 7 compara os resultados obtidos para a regressão sem o fator de expansão das observações e a nova regressão em que se incorpora o fator de expansão associado a cada pessoa na amostra. Esperava-se que no método de regressão por mínimos quadrados ponderados chegar-se-ia a resultados, na estimação dos coeficientes das variáveis, mais próximos dos valores obtidos por Ney e Hoffmann (2003): afinal de contas, esses autores alegam que a regressão feita por eles considera a referida expansão da observação com base no peso da pessoa. Contudo, o que se observa na tabela 7 é uma mudança não muito significativa na estimação dos coeficientes da variável pelo método com fator de expansão, em relação ao caso tradicional, que vemos adotando, em que não se considera o peso da observação na estimativa da equação de regressão.

Tabela 7 – Equações de rendimento estimadas para pessoas ocupadas na agricultura brasileira. PNAD 2001. Método dos mínimos quadrados ponderados. Com ou sem o fator de expansão associado a cada pessoa da amostra. Comparação com o resultado de Hoffman e Ney (2003).

	Coeficientes sem fator de expansão	Coeficientes com fator de expansão associado a cada pessoa na amostra	Hoffmann e Ney (2003)	Diferença em relação ao coeficientes das equações de rendimento estimadas por Hoffmann e Ney (2003)		
				Sem expansão	Com expansão	Aproximação com Hoffmann e Ney (2003)
intercepto	2,57463	123,8246	2,1498	19,76%	5659,82%	piorou
escolaridade	0,05881	0,07448	0,0603	-2,47%	23,52%	piorou
idade	0,35053	0,67874	0,314	11,63%	116,16%	piorou
idade2	-0,03521	-0,06925	-0,0326	-8,01%	-112,42%	piorou
sexo	-0,20219	-0,15199	-0,16	-25,20%	-5,89%	melhorou
contapp	-1,83839	-1,92843	-0,3702	*-396,59%*	-420,92%	piorou
empredor	-1,02569	-1,25976	0,3294	*-411,38%*	-482,44%	igual
preta	-0,03208	*0,01476*	-0,0841	61,85%	-117,55%	piorou
indígena	*-0,2658*	*0,27774*	0,0632	*-520,57%*	339,46%	melhorou
parda	-0,05045	*-0,01786*	-0,1036	51,30%	-82,76%	piorou
amarela	0,125	*0,61626*	0,5542	-77,44%	11,20%	melhorou
loghtrab	0,47085	0,79764	0,5125	-8,13%	55,64%	piorou
condunidom	0,19055	0,10042	0,1995	4,49%	-49,66%	piorou
NO	0,42115	*0,07125*	0,2369	77,78%	-69,92%	piorou
SE	0,29653	0,17632	0,1938	53,01%	-9,02%	melhorou
SP	0,6574	0,6328	0,5562	18,19%	13,77%	melhorou
SUL	0,51212	0,28099	0,3208	59,64%	-12,41%	melhorou
CO	0,54811	0,26037	0,4666	17,47%	-44,20%	piorou
domicilio	-0,0193	-0,03688	-0,1008	80,85%	63,41%	melhorou

logarea	0,15106	0,17749	0,1815	-16,77%	-2,21%	melhorou

Fonte: Elaboração do autor e Hoffmann e Ney (2003)

[1] Itálico indica não significativo em um nível de significância de 5%

Mudanças pronunciadas na estimativa do intercepto já eram esperadas, conforme já se argumentou; contudo, no exercício feito, os coeficientes estimados das variáveis pouco se alteraram. Mantiveram-se as discrepâncias em relação ao resultado obtido por Hoffmann e Ney (2003), reproduzidos na quarta coluna da tabela. Evidentemente a diferença nos coeficientes aumentou muito com o emprego do fator de expansão. Sobre a estimativa dos coeficientes, as distâncias dos correspondentes resultados desses autores aumentou com a regressão ponderada para as variáveis escolaridade, idade, idade2, contapp, preta, parda, loghtrab, condunidom, NO e CO. Para os outros coeficientes, o modelo com fator de expansão aproxima-se dos resultados de Hoffmann e Ney: a aproximação pode ser forte como em sexo, cor amarela, SE, SUL e logarea, ou tímida como em indígena, SP, domicílio e logarea.

A conclusão mais importante é a de que não foi, em nossa análise, a desconsideração do fator de expansão da observação (peso da pessoa) que nos afastou do resultado daqueles autores. Por exemplo, as acentuadas diferenças nos coeficientes estimados em condição do produtor (contapp e empredor) até piora com a expansão da observação pelo peso da pessoa. Então algum outro detalhe está afastando nossos resultados para esses coeficientes dos números obtidos pelos autores. É motivo de preocupação neste estudo o fato de não termos replicado perfeitamente os resultados obtidos pelo estudo de comparação que estamos utilizando. As informações do método de tratamento dos dados de Hoffman e Ney elucidariam a razão destas diferenças.

Desafios futuros da pesquisa e os possíveis desdobramentos

Há um conjunto de questões que ficaram pendentes e que se deve enfrentar. Referimo-nos especialmente à interpretação e aprofundamento do tratamento empírico das PNADs rurais. Já antecipamos aqui a necessidade de estudar-se a razão das disparidades apresentadas para a PNAD de 2001 em relação ao estudo de referência de Hoffman e Ney (2003). Outro desafio é o de avançar-se no estudo do significado dos coeficientes de regressão das variáveis que obtivemos em nosso tratamento estatístico dos dados. Um estudo que se fará necessário é o da chamada contribuição marginal de cada fator para a soma de quadrados de regressão. Com esses indicadores, poderemos avaliar a importância relativa da variável em cada modelo sem nos limitarmos apenas à consideração dos coeficientes de regressão, conforme visto neste ensaio.

Bibliografia Consultada

ALCHIAN, A A. “Some Economics of Property Rights.” *Il Politico*, 30 (No. 4): 816-829, 1965.

ÁVILA, Antonio F. D., Evenson, Robert E., “Crescimento da Produtividade Total dos Fatores: o papel do capital tecnológico”, *Revista de Política Agrícola,* XIV (No. 2), abril a junho, Brasília, 2005.

BALSADI, Otavio V., Comportamento das ocupações na agropecuária Brasileira no período 1999-2003, *Informações Econômicas*, 35 (No. 9), setembro, São Paulo, 2005.

CALABRESI, G. "Some Thoughts on Risk Distribution and the Law of Torts." *Tale Law Journal*, 70 (No. 4, March): 499-553, 1961.

COASE, R. H., *The Firm, the Market and the Law*, The University of Chicago Press, 1988, pp. 33-55; in Williamson, O. and Winter, S., eds. (op. cit.).

DEMSETZ, H. "The Exchange and Enforcement of Property Rights." *Journal of Law and Economics* 3 (October): 1-47, 1964.

EGGESTSSON, Thráinn, Economic Behavior and Institutions, Cambridge Univ. Press, Cambridge, 1999.

FISHLOW, Albert, "Distribuição de renda no Brasil: um novo exame", Dados (No. 11), Rio de Janeiro, 1973.

GASQUES, José G., Bastos, E. T., Bacchi. M. P. R., Conceição, J. C. P. R. da, Condicionantes da Produtividade da Agropecuária Brasileira. *Revista de Política Agrícola*, ano XIII, (No. 3): 73-90, julho a setembro, Brasília, 2004.

HOFF, Karla (2000), "Distribuição de Riquezas, Eficiência Econômica e Incentivos: uma perspectiva da economia da informação". Em: Edson Teófilo (org.), Distribuição de Riqueza e Crescimento Econômico, *Núcleo de Estudos Agrários e Desenvolvimento Rural Sustentável*, Ministério do Desenvolvimento Agrário, Brasília, 2000.

HOFFMAN, Rodolfo, "Equações de rendimento para Pessoas Ocupadas no Brasil: contrastes regionais e setoriais". Anais do XXXVI Congresso Brasileiro de Economia e Sociologia Rural, Poços de Caldas 1998.

HOFFMAN, Rodolfo, Ney, M. G., "Desigualdade de Renda na Agricultura: oi efeito da posse da terra", *Economia*, 4 (No. 1), janeiro a junho, Niterói (RJ), 2003.

INTERNATIONAL SERVICE FOR NATIONAL AGRICULTURAL RESEARCH, ISNAR Annual Report, The Netherlands (s.d.) http://www.isnar.cgiar.org/about_isnar/ar/index.htm

JENSEN, Michael C., Meckling, W. H., "Rights and Production Functions: An Application to Labor-Managed Firms and Codetermination." *Journal of Business* 52 (No. 4): 469-506, 1979.

LANGONI, C; G. "Distribuição da renda: resumo da evidência", *Dados* (No. 11), Rio de Janeiro, 1973.

MADDALA, G. S., Econometrics, McGraw-Hill, Economics Series, 1977.

NELSON, R., Winter, S. An Evolutionary Theory of Economic Change. Cambridge, Mass.: Harvard University Press, 1982.

NORTH, Douglass, Structure and Change in Economic History. New York: W.W. Norton and Company, 1981.

PEJOVICH, Steve, "Karl Marx, Property Rights School and the Process of Social Change." *Kyklos*, 35 (No. 3): 383-397, 1982.

ROSS, Stephen, "The Economic Theory of Agency: The Principal´s Problem." *American Economic Review* 62, May: 134:139, 1973.

SHEAHAN, John, Patterns of Development in Latin America: poverty, repression, and economic strategy, Princeton University Press, 1987.

VEIGA, José E. da, “Pobreza Rural, Distribuição da Riqueza e Crescimento: a experiência Brasileira”, Em: Edson Teófilo (org.), Distribuição de Riqueza e Crescimento Econômico, *Núcleo de Estudos Agrários e Desenvolvimento Rural Sustentável*, Ministério do Desenvolvimento Agrário, Brasília, 2000.

WILLIAMSON, Oliver E. The Economic Institutions of Capitalism, New York, The Free Press, 1995.

WILLIAMSON, Oliver E.; Winter, Sidney G. The Nature of the Firm: origins, evolution, and development, Oxford, Oxford University Press, 1993.