Rural Poverty in Brazil and the Role of the Non-agricultural Sector*

ERIK JONASSON**
Lund University, Lund, Sweden

and

STEVEN M. HELFAND
University of California, Riverside, USA

Draft, June 1, 2007

ABSTRACT

Thirty percent of the rural labor force in Brazil has non-agricultural jobs as principal employment. With average earnings being higher in the rural non-agricultural sector than in agriculture, rural non-agricultural employment (RNAE) receives increasing attention as a potential pathway out of rural poverty. To assess for whom and under what circumstances RNAE is a viable source of income, this paper analyzes the determinants of non-agricultural employment and earnings potential in non-agricultural jobs. Education stands out as one of the key determinants of employment outcome as well as of earnings potential. Failure to control for demand side effects, however, might cause over-estimation of individual and household-specific characteristics. The empirical results show that local market size, rural infrastructure, and distance to population centers have a major impact on non-agricultural employment prospects.

I. INTRODUCTION

Due to large-scale rural-to-urban migration there are now more urban poor than rural poor in Brazil. In relative terms, however, poverty is more evident in rural than in urban areas. The rural poverty rate is 61 percent, more than twice as high as the urban poverty rate (Demographic Census, 2000). The vast

* This material is based upon work supported by a grant from the United States Agency for International Development (AID), and a subcontract from Broadening Access and Strengthening Input Market Systems (BASIS)/ Collaborative Research Support Program (CRSP)/University of Wisconsin - Madison awarded to the Regents of the University of California, Riverside. The opinions, findings, conclusions, and recommendations are solely those of the authors.

** Department of Economics, Lund University; P.O. Box 7082; SE-22007 Lund; Sweden.
erik.jonasson@nek.lu.se

majority of the rural poor lives in remote areas and has peasant farming or agricultural labor as principal means of income. With this character of rural poverty, the main component of the World Bank's latest comprehensive strategy proposal to fight rural poverty in Brazil is to strengthen the small-farm sector. Parallel, The World Bank (2003) recommends migration of the young to more economically developed urbanized areas, and stimulation of growth of the rural non-agricultural sector. The Bank's general conclusion regarding the rural non-agricultural (RNA) sector is that it is unlikely to constitute a potential poverty exit path for the bulk of the rural poor. On the one hand, most of the rural poor reside in remote rural areas; on the other hand the viability of the RNA sector depends on proximity to input and output markets and a certain level of local infrastructure. Nonetheless, 28 percent of the rural population has rural non-agricultural employment (RNAE) as their principal occupation – and among those who participate in the RNA sector, poverty is lower. Even though the share of RNAE is lower in Brazil than the 40-percent average for Latin America, it is evident that non-agricultural activities take place far beyond the urban periphery.

This paper aims to assess who participates in the rural non-agricultural sector and what determines the earnings potential of workers in this sector. Empirical studies suggest that a combination of household-specific and locational characteristics determine the rural household's probability to engage in RNAE and its non-agricultural earnings potential. Usually, education stands out as the key household or individual asset for participation in the RNA sector. In terms of extra-household determinants, proximity to markets and the level of infrastructural development correlate positively with income and the degree to which households engage in RNAE. Household asset endowments on their own will not generate upward income mobility if market participation is very costly due to physical distance to markets and underdeveloped infrastructure that obstruct the mobility of people, capital, goods, and information.

Even though there is a consensus in the literature that location does matter for the viability of the RNA sector, the empirical support so far relies on indirect locational indicators, which give us limited insight into the role that remoteness from markets and urban areas actually play. Failure to control for local characteristics may over-estimate the role of individual or household-specific characteristics when drawing inference about rural household labor allocation and livelihood strategies. Besides controlling for broad geographical differences, with, for example, macro regional dummies, authors have used various indicators to capture local economic conditions. In studies that focus on the Brazilian North-East and South-East, Ferreira and Lanjouw (2001) and Lanjouw (2003) distinguish between five degrees of rurality and find that households residing in urban extensions have a higher probability than households in rural-exclusive areas to engage in RNAE. More than 85 percent of the rural population, however, falls into the rural-exclusive category of location, and we know little about the role of location for this part of the rural labor force. Corral and Reardon (2001) find that distance to the nearest health center correlates negatively with RNAE in Nicaragua. Lanjouw (2001) uses distance to the nearest secondary school in rural Ecuador as a proxy for distance to nearest town or settlement but finds no statistically significant relationship with RNAE. Isgut (2004), in a study on rural non-agricultural income determinants in Honduras, uses local averages of household characteristics to estimate regional effects. He finds that the degree of urbanization increases household non-farm income shares, while local average income correlates negatively with non-farm income. De Janvry and Sadoulet (2001) test to what extent the number of urban and rural centers within one hour's commuting distance matters for RNAE in Mexico, but find a statistically significant effect only for female RNA wage laborers.

In the assessment of determinants of rural non-agricultural employment and income, our study goes beyond categorical controls for location. By utilizing data from the Demographic Census of 2000, we have been able to include disaggregated variables to test for the role of municipal-level economic factors, such as local market size and distance to major population centers. Our results show that, in line with previous research, personal as well as household characteristics do matter for employment outcomes as well as for income earnings potential. Education strongly increases the probability that a rural worker is engaged in RNAE, particularly the probability of being engaged in high-productivity RNAE. Demand-side factors, however, have strong effects too on both RNAE and earnings. Not only do the degree of urbanization and the level of rural infrastructure increase RNAE opportunities. Market size and the distance to population centers also have large effects on employment outcomes. The demand-side factors, however, do not render individual characteristics insignificant.

II. POVERTY AND NON-AGRICULTURAL EMPLOYMENT IN RURAL BRAZIL

The Brazilian population amounted to 169.8 million in the year 2000 (Demographic Census, 2000). Less than 32 million, or 19 percent, of the population are considered rural residents. Whereas this share is close to the average for Latin America, it is much lower than other developing regions such as South Asia (72 percent) and Sub-Saharan Africa (64 percent). Table 1 gives an overview of the Brazilian population and of poverty rates across regions. The North and North-East are the least urbanized regions, in which almost a third of the population live in rural areas. In the densely populated South-East, only 9 percent of the households are rural. The Census classifies rural households into five rural sub-categories. We refer to these as 1) rural agglomerations that are urban extensions (*extensão urbana*), 2) isolated rural agglomerations or towns that have some service provision (*povoado*), 3) isolated rural agglomerations linked to a single landowner (*núcleo*), 4) other isolated agglomerations (*outors agglomerados*), and 5) rural areas exclusive of agglomerations (*zona rural exclusive aglomerado rural* – rural areas that do not qualify for any of the definitions of agglomerations above). The vast majority, 86 percent of the rural population, fall into the latter category. Only 11 percent live in what is referred to as rural towns or agglomerations and most of the remaining 3 percent live in areas considered as urban extensions.

Rural poverty

The high degree of urbanization in Brazil makes poverty, in absolute terms, primarily an urban problem. In relative terms, however, poverty is most prevalent and most severe in rural areas. Rural poverty for Brazil as a whole is estimated at 61 percent. The poorest region is the North-East, in which 77 percent of the rural population is classified as poor. This is also the region where almost half of Brazil's rural population resides. Rural poverty is lower in the South, South-East, and Center-West regions, where it ranges between 35 and 43 percent. Urban poverty is 25 percent nationwide and, again, highest in the North-East. The poverty headcount ratio and the poverty gap reported in this paper are based on a poverty line set to 75 Reais per month, which corresponds to half the minimum wage of year 2000 (this poverty line is also used by the Atlas of Human Development, IPEA/UNDP). Since the income measure used here refers to *monetary* income reported in the Census, it under-estimates the total income for farm households, which use part of the farm production for consumption. Hence, even if monetary income approximates fairly well the total income of urban and rural non-agricultural households, the under-estimation of farm household incomes inflates to a certain degree the rural poverty rates.[1] Based on a comparison with the Census of 1991, rural poverty fell by 11 percentage points during the 1990s (Helfand and Levine, 2004). With the exception of the North, poverty fell in all regions and mostly so in the South and South-East regions. Based on the yearly PNAD survey, The World Bank (2006) shows that rural poverty fell by as much as 10 percentage points nationwide during 1995 to 2004.

[1] The poverty rates reported here are similar to income-based poverty lines reported by OECD (2005), based on the Census, and by the World Bank (2006), based on the 1996 PNAD survey. They are considerably higher, however, than the expenditure-based poverty rates reported by Lanjouw (2003).For a detailed analysis of income- versus expenditure-based poverty measures, see Figueiredo, Helfand, and Levine (2007).

TABLE 1: Population and poverty

	Brazil	*North*	*North-East*	*South-East*	*South*	*Center-West*
population, total (millions)	169.8	12.9	47.7	72.4	25.1	11.6
urban	138.0	9.0	33.0	65.5	20.3	10.1
rural	31.8	3.9	14.8	6.9	4.8	1.5
population share, rural	0.19	0.30	0.31	0.09	0.19	0.13
Rural population shares by rural sub-category						
Rural agglomerations that are urban extensions	0.03	0.02	0.02	0.10	0.02	0.01
Isolated rural agglomerations or towns	0.11	0.12	0.16	0.05	0.02	0.06
Isolated rural agglomerations or clusters	<0.01	0.01	0.01	<0.01	<0.01	<0.01
Other isolated agglomerations	<0.01	<0.01	<0.01	0.01	<0.01	0.01
Rural areas exclusive of agglomerations	0.86	0.85	0.82	0.85	0.96	0.92
Poverty, headcount ratio (P0)	0.32	0.48	0.55	0.19	0.19	0.24
Urban	0.25	0.39	0.45	0.16	0.16	0.21
Rural	0.61	0.70	0.77	0.42	0.35	0.43
Urban extensions	0.42	0.43	0.57	0.24	0.26	0.30
Isolated rural agglomerations or towns	0.58	0.66	0.72	0.43	0.33	0.42
Rural areas exclusive of agglomerations	0.62	0.72	0.79	0.45	0.36	0.44
Poverty gap (P1)	0.49	0.54	0.56	0.47	0.44	0.44
urban	0.47	0.48	0.50	0.47	0.43	0.43
rural	0.56	0.60	0.62	0.48	0.46	0.48
Urban extensions	0.50	0.47	0.52	0.50	0.48	0.45
Isolated rural agglomerations or towns	0.55	0.58	0.59	0.50	0.46	0.48
Rural areas exclusive of agglomerations	0.57	0.61	0.63	0.48	0.46	0.48

Source: Demographic Census 2000, authors' calculation.

There is a clear positive correlation between poverty and degree of rurality. Within each region, poverty is lowest in the urban areas and increases by rural sub-category. The poverty gap, which is an indicator 'how poor' the poor are, follows the same geographical pattern as the poverty headcount ratio. Where poverty is most prevalent, people also tend to fall farthest beneath the poverty line. In the rural North-East, for example, the monetary income of the average poor is 62 percent below the poverty line.

Rural employment

Around 70 percent of the rural labor force have their principal employment in agriculture (cultivation, animal rearing, and forestry). The remaining 30 percent are employed in the rural non-agricultural sector. As shown in Table 2, there are regional variations in the composition of the rural labor force. The North-East is not only the poorest region, but also the region with the highest share of the rural labor force engaged in cultivation (66 percent) and with the lowest share in the RNA sector

(25 percent). Rural non-agricultural employment is highest in the relatively densely populated and highly urbanized South-East region. Rural areas in the urban-extension category are largely dominated by non-agricultural work; only 10 percent of the labor force in these areas are involved in agriculture. Also in rural towns, non-agriculture activities dominate agriculture in terms of employment.

Figures A1–A5 in the appendix give a picture of the structure of the rural labor force and how non-agricultural employment extends outside urban regions. The figures are maps of each of the five macro regions of Brazil and depict the share of the rural labor force whose principal occupation is in RNAE in each municipality. It is evident that non-agricultural activity is more prevalent in the proximity of major cities and urbanized areas. The North-East, South-East, and South show a concentration of non-agricultural activities near urban regions along the coast line. The pattern is most pronounced in the densely populated areas surrounding São Paulo and Rio de Janeiro. Given, however, that rural households close to urbanized areas are situated within commuting distance, we tend to overstate the extent of rural non-agricultural jobs available to most rural residents. In fact, if we exclude rural residents who live within 50 kilometers distance to a population center of 100,000 people or more, the share of RNAE in the rural labor force decreases from 30 to 24 percent. In the South-East, RNAE would fall from 39 percent to 25 percent.

TABLE 2: Percentage of rural labor force by sector of principal occupation

	Cultivation	Animal rearing	Forestry	Non-agriculture	*Total*
Region					
Brazil	0.56	0.12	0.02	0.30	*1.00*
North	0.52	0.12	0.04	0.32	*1.00*
North-East	0.66	0.07	0.03	0.25	*1.00*
South-East	0.43	0.16	0.01	0.39	*1.00*
South	0.56	0.15	0.02	0.27	*1.00*
Center-West	0.27	0.41	0.02	0.30	*1.00*
Rural sub-category					
Urban extension	0.08	0.02	0.00	0.90	*1.00*
Rural towns	0.38	0.06	0.02	0.54	*1.00*
Rural exclusive	0.60	0.13	0.02	0.25	*1.00*
Employment status					
wage labor	0.31	0.15	0.02	0.52	*1.00*
self-employed	0.60	0.11	0.03	0.26	*1.00*
non-paid	0.83	0.10	0.02	0.05	*1.00*
Gender					
men	0.59	0.14	0.02	0.25	*1.00*
women	0.48	0.07	0.03	0.42	*1.00*

Source: Demographic Census 2000, authors' calculation.

Most people in the rural labor force can be divided into three broad categories: wage laborers, self-employed, and unpaid workers (working either as unpaid household members, trainees, or in subsistence agriculture). Each of these groups constitutes about a third of the rural labor force. Half of the wage laborers are engaged in RNAE, a fourth of the self-employed, and only a very small share of the unpaid workers. The group of unpaid workers consists almost exclusively of farm workers. This implies that the majority of those who work in the RNA sector are wage laborers; 68 percent of RNAE takes the form of wage labor, 27 percent is self-employment, and the rest non-remunerated RNAE. The share of RNAE is considerably higher among women (42 percent) than men (25 percent).

As a residual concept, the rural non-agricultural sector contains a wide range of activities, including everything from low-return street-vending to well-paid jobs in the formal sector. Table 3 shows that the three largest RNA sectors – manufacturing, commerce, and domestic services – employ almost 50 percent of the non-agricultural labor force. Education and construction constitute another one-fifth of the non-agricultural labor force. Manufacturing employs a considerably higher share (30 percent) in the North and South than in the other regions. Domestic services play a larger role in South-East and Center-West than in other regions (21 and 23 percent of RNAE, respectively). In the North, fishing is the second largest RNA sector in terms of employment, whereas it plays a minor role in other regions. Among wage laborers, domestic services are the largest sector of non-agricultural employment, and most workers in this sector are women. Among self-employed engaged in non-agricultural activities, manufacturing and commerce are the two major sectors. The most noticeable difference between male and female non-agricultural work is that women are more represented in domestic services and education than men. Men are to a higher extent engaged in traditional male-dominated jobs such as construction, transportation, and fishing.

On average, people earn higher incomes in the rural non-agricultural sectors than in agriculture. This is true both for laborers and self-employed as well as for both men and women. Table 4 shows average monthly earnings in the five non-agricultural sectors that employ the majority of the RNA labor force. The average earnings in agriculture is R$280 (when considering earned monetary income from principal employment and excluding those with reported zero income).[2] Domestic services are the only major RNA sector in which average earnings are lower. Self-employed earn more than laborers, and in all sectors men earn more than women.

[2] Whereas few non-agricultural workers report zero earned income, more than a third of the agricultural labor force does so. Given that the estimated mean of earned income reported in Table 4 includes only those with positive earned income, it most likely overestimates the actual earnings of the average agricultural worker.

TABLE 3: Rural non-agricultural employment (RNAE) by sub-sector (% with primary occupation)

	Region						*Employment*		*Gender*	
	Brazil	North	North-East	South-East	South	Center-West	Labor	Self-employed	Men	Women
Manufacturing	0.20	0.25	0.18	0.18	0.29	0.16	0.18	0.22	0.23	0.17
Commerce	0.14	0.13	0.14	0.15	0.15	0.15	0.09	0.27	0.17	0.10
Domestic Services	0.14	0.08	0.12	0.21	0.13	0.23	0.21	0.00	0.05	0.28
Education	0.11	0.10	0.14	0.06	0.07	0.11	0.16	0.01	0.03	0.22
Construction	0.10	0.05	0.11	0.12	0.09	0.07	0.10	0.12	0.16	0.00
Public administration	0.06	0.05	0.07	0.04	0.05	0.06	0.09	0.00	0.05	0.07
Hotel and restaurants	0.05	0.04	0.05	0.06	0.04	0.07	0.03	0.10	0.05	0.05
Fishing	0.05	0.18	0.05	0.01	0.01	0.02	0.01	0.09	0.07	0.01
Transportation	0.05	0.03	0.05	0.05	0.06	0.04	0.04	0.08	0.08	0.00
Real estate services	0.03	0.02	0.03	0.04	0.03	0.03	0.03	0.04	0.04	0.02
Other social and personal services	0.03	0.02	0.03	0.03	0.03	0.03	0.02	0.04	0.02	0.03
Extraction	0.02	0.03	0.01	0.01	0.02	0.02	0.02	0.02	0.03	0.00
Health care and social services	0.01	0.01	0.01	0.02	0.01	0.01	0.02	<0.01	0.01	0.03
Utilities	<0.01	<0.01	<0.01	0.01	<0.01	<0.01	0.01	<0.01	0.01	<0.01
Financial services	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Total	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*

Source: Demographic Census 2000, authors' calculation.

Even though average earnings in most of the RNA sectors are higher than in agriculture, there are also many low-paid non-agricultural jobs. We divided RNAE into two groups depending on earnings relative to agriculture. If an individual is engaged in RNAE and has earnings below the municipal average agricultural wage rate, we considered the individual as being engaged in low-productivity RNAE, and those who earn above this average agricultural wage rate, as being engaged in high-productivity RNAE. With this categorization, 53 percent of the RNAE labor force is considered high-productivity RNAE. As indicated by the last column in Table 4, in the educational sector, more than two-thirds have high-productivity jobs, but only one-fifth of those occupied with domestic services.

Non-agricultural activities are sometimes considered primarily as a means of income diversification among rural households (Barrett et al., 2001; Ellis, 2000). For households in rural Brazil, however, income diversification per se does not appear to be a deliberate strategy of the majority of households. We defined households as agriculture specialized if they derive 90 percent or more of their earned income from agriculture; non-agriculture specialized if they derive 90 percent or more from RNAE; and pluriactive otherwise. Only 14 percent of the rural households are considered pluriactive by this definition. As seen in Table 5, this pattern is confirmed across all regions.

TABLE 4: Rural non-agricultural income by sector (R$ per month, 2000)

Sector	*Brazil*	*Labor*	*Self-employed*	*Men*	*Women*	*Share high productivity*
Manufacturing	337	314	385	390	209	0.51
Commerce	449	310	578	492	329	0.57
Domestic Services	160	160	n/a	223	140	0.21
Education	295	292	411	394	274	0.68
Construction	334	299	402	335	321	0.65
Public administration	387	387	n/a	507	256	0.64
All RNA sectors	346	294	479	416	236	0.53
Compare to: agricultural sectors	280	198	346	296	170	n/a

Note: Share high-productive refers to the fraction of workers in the RNA sector who are engaged in high-productivity RNAE. The exchange rate US$/R$, July 2000, was 0.55. *Source*: Demographic Census 2000, authors' calculation.

TABLE 5: Rural household specialization (by source of earned income)

	Brazil	*North*	*North-East*	*South-East*	*South*	*Center-West*
Agricultural specialization	0.57	0.62	0.58	0.50	0.59	0.63
Pluriactive household	0.14	0.12	0.14	0.15	0.15	0.14
Non-agricultural specialization	0.29	0.26	0.28	0.35	0.25	0.23
Non-agr spec by income group:						
Lowest income quintile	0.21	0.21	0.18	0.30	0.16	0.24
2nd	0.23	0.22	0.20	0.32	0.23	0.21
3rd	0.27	0.25	0.25	0.34	0.27	0.21
4th	0.30	0.26	0.30	0.36	0.29	0.22
Highest income quintile	0.37	0.33	0.38	0.42	0.30	0.26
Poor	0.24	0.24	0.24	0.26	0.19	0.21
Non-poor	0.36	0.31	0.38	0.41	0.29	0.24

Source: Demographic Census 2000, authors' calculation.

Noticeable in terms of specialization is that richer households are to a larger extent engaged in RNAE than poorer households. In the lowest income quintile of rural households, 21 percent of the households are specialized in non-agriculture, whereas the share is almost twice as high (37 percent) in the highest income quintile. The positive correlation between household income and RNAE is consistent with several other country studies in Latin America (Ecuador – Lanjouw, 1998; Mexico – de Janvry and Sadoulet, 2001; Peru – Escobal, 2001). Some studies have suggested a U-shaped relationship between household income and the degree of non-farm activity, with RNAE shares highest among the richest and the poorest households and lowest among the average income households (Ellis, 2000). The rationale for this sometimes observed pattern is that the poor rural households seek non-agricultural employment based on a "push-factor" of insufficient income from own farm production. For the richer households, on the other hand, RNAE shares are high as a result of the "pull-factor" in terms of higher returns to education and capital and access to the well-paid rural jobs.

Differences in average earnings suggest that the rural non-agricultural sector can indeed offer a potential pathway out of poverty. This is not necessarily the case, however, if supply-side factors play an important role alongside individual characteristics for non-agricultural opportunities in rural areas. In the next two sections, we analyze the importance of these supply and demand side effects, firstly by assessing what influences the probability that people in the rural labor force engage in non-agricultural activities, and secondly what determines the earnings of those engaged in non-agricultural activities.

III. DETERMINANTS OF RURAL NON-AGRICULTURAL EMPLOYMENT

In this section we report the results of a probability analysis of engagement in rural non-agricultural employment. The main data source used for the analysis is the Demographic Census of year 2000. We also used several variables intended to capture local characteristics, such as potential NAE opportunities or distance to population centers.[3] The analysis focuses firstly on the probability that a rural worker is engaged in any kind of RNAE. Due to the heterogeneity of the RNA sector, we then distinguish between what we defined in the previous section as 'high-productivity' and 'low-productivity' non-agricultural activities, and study separately the probabilities of engaging in these two RNAE types.

Estimation method

We estimated the individual's probability of engaging in RNAE by three approaches. In the first approach, a binomial probit model was estimated, in which the dependent variable is the binary *RNAE* variable, indicating whether the individual is engaged in RNAE as opposed to agriculture. In the second approach, we estimated separate probit models to assess what determines the selection process into high- and low-productivity RNAE. In the third approach, we estimated a multinomial probit model, in which we analyzed jointly the probabilities of engaging in agriculture and high- and low-productivity RNAE.

[3] We thank Eustáquio Reis, Marcia Pimentel, and the Applied Economic Research Institute (IPEA) for assistance with the construction of these variables.

The binomial model is specified based on the assumption that a set of exogenous variables determines an endogenous, but unobserved (latent), variable V_i. If V_i exceeds a certain threshold value, V_i*, the individual is engaged in RNAE; otherwise, he or she is engaged in agriculture. The latent variable V can be thought of, in this case, as the rural worker's expected earnings if participating in the rural non-agricultural sector. The threshold V* could be the shadow wage for agricultural work on the own farm or the wage rate on the agricultural wage labor market.[4] Thus, the probability of individual i being engaged in RNAE, P_i, is modeled as the probability that V_i exceeds V_i*:

$$P_i = PROB_i\left(RNAE_i = 1 \middle| X_{ijk}, H_{jk}, M_k\right) = PROB(V_i \geq V_i^*) \quad (1)$$

in which X_{ijk}, H_{jk}, and M_k denote vectors of variables to characterize, respectively, individual i, household j to which the individual belongs, and municipality k in which the households is situated. Let v_i denote the difference V_i–V_i*, which is the expected net gain from RNAE. This net gain is modeled as a log-linear function of X, H, and M:

$$v_i = X_{ijk}\beta_1 + H_{jk}\beta_2 + M_k\beta_3 + \varepsilon_{ijk} \quad (1')$$

where the βs are row vectors of coefficients to be estimated, and ε is a residual assumed to be normally distributed with zero mean and variance σ^2. Let F(.) be the standard normal cumulative distribution function of ε; then individual i's probability of engaging in RNAE can be estimated as:

$$P_i = PROB\left(X_{ijk}\beta_1 + H_{jk}\beta_2 + M_k\beta_3 \geq -\varepsilon_{ijk}\right) = F\left(X_{ijk}\beta_1 + H_{jk}\beta_2 + M_k\beta_3\right) \quad (1'')$$

In the second approach, we applied the distinction between low- and high-productivity RNAE that we introduced in the previous section (letting the average agricultural wage rate in the municipality be the benchmark separating the two). Two binomial probit models were estimated along the same lines as the model above, assessing separately the determinants of the probability of being engaged in the two types of RNAE:

$$P_i^{LOW} = PROB_i\left(RNAE_i^{LOW} = 1 \middle| X_{ijk}, H_{jk}, M_k\right) = F\left(X_{ijk}\beta_1 + H_{jk}\beta_2 + M_k\beta_3\right) \quad (2)$$

$$P_i^{HIGH} = PROB_i\left(RNAE_i^{HIGH} = 1 \middle| X_{ijk}, H_{jk}, M_k\right) = F\left(X_{ijk}\beta_1 + H_{jk}\beta_2 + M_k\beta_3\right) \quad (2')$$

[4] In a broader sense, V could also be interpreted as a subjective utility measure of the individual, so that RNAE is chosen if the expected utility of RNAE is higher than the expected utility (V*) of agricultural work.

where the *LOW* and *HIGH* superscripts distinguishes the two RNAE forms. If $RNAE^{LOW}$ is zero, the individual either works in agriculture *or* in high-productivity RNAE. Similarly, if $RNAE^{HIGH}$ is zero, the individual either works in agriculture or in low-productivity RNAE.

The $RNAE^{LOW}$ model creates somewhat odd comparison employment categories (agriculture or high-productivity RNAE). Therefore, we complemented these models with a multinomial probit model, in which the probabilities of engaging in low- and high-productivity RNAE are analyzed jointly with the probability of being involved in agricultural work. The third model is specified as:

$$P_i^e = PROB_i\left(EMP_i = e \middle| X_{ijk}, H_{jk}, M_k\right) = F\left(X_{ijk}\beta_1^e + H_{jk}\beta_2^e + M_k\beta_3^e\right) \qquad (3)$$

in which P^e is the probability that individual *i* has employment *e*; *e* being either i) agricultural work, ii) low-productivity RNAE, or iii) high-productivity RNAE. As in the previous probability models, P^e is modeled as a function of the log-linear combination of *X*, *H*, and *M* and its corresponding coefficients. The difference in this model compared with models (1) and (2) is that the estimated coefficients vary by employment category.

Variables and data source

The Demographic Census data include more than 20 million observations (12 percent of the population) and are constructed to be representative at municipal level. There are 5,507 municipalities defined for the 2000 Census, with an average population of a little more than 30,000 people; a few have over a million. In the empirical analysis, we used the rural adult labor force as the base sample, which includes 1.76 million individual observations. Adults were defined as everyone 15 years or older. Anyone reporting an occupation was considered as a participant in the labor force, including unpaid workers. Table A1 in the appendix provides an overview of the Census data and how the base sample was selected for the empirical analysis. Descriptive statistics of the variables included in the regression analysis are provided in Table 6.

The endogenous RNAE variable used as in the models to identify rural non-agricultural employment was based on reported principal occupation. Turning to the explanatory variables, the individual characteristics included in vector *X* include age, gender, race, education, migrant-status, and an indicator for macro-region. Age, age-squared, and years of schooling serve as proxies for human capital. Even though human capital matters for agricultural labor productivity, the non-agricultural sector is likely to contain those jobs with the highest returns to education, and would hence attract the relatively well-educated workers in the rural labor force. Human capital can also have the allocative effect of allowing households to make the optimal labor allocation decision (Yang and An, 2002; Laszlo, 2005). Education was controlled for by three dichotomous variables, which are based on the number of completed years of schooling. Zero education is the benchmark category and contains 24 percent of the rural labor force. Gender was included to control for systematic differences between male and female workers in terms of job preferences, work hours, but also demand-side effects, such as gender discrimination in payment schemes. Race, controlling for black and indigenous, was included for similar reasons. A dummy variable was included, indicating whether the individual has always lived in the municipality, and is a proxy for migrant-status. The World Bank (2003) notes that income differentials are the single largest driving force to explain migration. People who have moved could have a lower opportunity cost of staying on the farm and would be more willing to seek non-agricultural employment. Migration could also be an indicator of mobility and hence ability to engage in the employment with highest returns for the individual.

Household characteristics, *H*, include the number of household members, average education in the household (excluding individual *i*'s education), and an indicator variable for electric lighting. The size of the household was included to control for opportunities to employment diversifications: the larger the labor supply in the household, the larger the opportunities to devote some household labor to non-agricultural activities. Average years of schooling among other household members is a proxy for the household stock of human capital. Given that there are some spillover effects within the household, the higher the average education, the more likely it is that an individual undertakes employment with skill requirements. Electric lighting is a proxy for local infrastructure development in the immediate neighborhood of the household, which is assumed to correlate positively with RNAE opportunities. Two variables were included that indicate whether the household is situated in a rural town or urban extension, as opposed to a rural exclusive area (cf. Table 2).

The municipal-level characteristics, *M*, included in the model are of three types: 1) indicators for rural infrastructure and degree of urbanization in the municipality, 2) measures of local market size, and 3) measures of distance to the nearest population center with a certain population size. Their descriptive statistics are found in the lower part of Table 6. These variables were used in alternative model specifications to assess the role of the local economic geography for rural non-agricultural activities. Whereas individual and household characteristics can be regarded as supply-side factors, the municipal-level characteristics can be considered demand-side factors (demand for non-agricultural goods and services, and hence demand of RNA labor). The indicators of rural infrastructure development include share of rural households in the municipality with access to electricity and telephone. Whereas the average share of rural households with electricity is 74 percent, the share of households with a telephone line is only six percent. These six percent largely represent those in the category residing in 'rural towns'. The degree of urbanization is given by the share of households in the municipality that are classified as urban. These demand-side variables can be considered as indicators of market participation cost: The lower the degree of urbanization and infrastructural development, the higher the cost of participation in the input and output markets, and the lower the prospects for non-agricultural activities.

The measures of local market size, the second group of municipal-level variables, include population in municipality and distance-weighted measures of regional population. The municipal population (*Mun_pop*) gives an idea of the size of immediate local market size. To get a more accurate proxy of the actual local market size, we used two different distance-weighted municipal-level population measures. The first one (*Pop1d*) was defined by adding to the municipal population the population of other municipalities, weighted by the inverse kilometer distance to each municipality. The second population measure (*Pop100d*) uses a linearly declining weight, which takes into account only population within a 100-kilometer distance (weight=1 at zero km, weight=0 at 100km). Formally,

$$Pop1d_k = Mun\_pop_k + \sum_{l \neq k} Mun\_pop_l \left(1 / D_l\right) \quad (4)$$

$$Pop100d_k = Mun\_pop_k + \sum_{l \neq k} Mun\_pop_l \left(1 - D_l / 100\right) \quad (4')$$

where D_l is the distance in kilometers to the seat of municipality l . As seen by the large difference in means between the two, *Pop1d* discounts much more heavily for distance than *Pop100d* (e.g. population in a municipality 50 kilometers away only gets a 2% weight with *Pop1d,* but a 50% weight with *Pop100d*). They both have the advantage of taking into account people outside the own municipality, as a decreasing function of distance.

The measures of distance to population centers, thirdly, were included as an alternative proxy for regional demand and access to markets. Distances were estimated to the nearest municipality with 50–100, 100–250, 250–500, and more than 500 thousand people, respectively. They are defined as the minimum distance between the seat of the municipality and the seat of the nearest municipality with the respective population size. To adjust for the average distance from a farm to the municipal seat in the location of origin, we added a distance measure as a function of the area of the municipality. Distance to the own municipal seat, d_k, was estimated by assuming that the municipality was circle shaped, with the municipal seat in the middle, and with the average rural household being situated half a radius distance to the seat, such that $d_k = 0.5\sqrt{A / \pi}$, where A is area of municipality in square-kilometers.

TABLE 6: Summary statistics of variables used in the empirical analysis

Variable	*Mean*	*Std. Dev.*	*Min*	*Max*	*Description*
RNAE	0.28	0.45	0.00	1.00	Has RNAE as principal employment (d)
RNAE low	0.14	0.35	0.00	1.00	Low-productivity RNAE
RNAE high	0.14	0.35	0.00	1.00	High-productivity RNAE
Y^{NA}	334	1,170	1	350,000	Earned non-agricultural income for Y^{NA} >0
Y^{AGR}	282	1,331	1	500,000	Earned agricultural income, for Y^{AGR} >0
Age	36	15	15	100	Individual's age
Age2	1,526	1,217	225	10,000	Individual's age, squared
Male	0.71	0.45	0.00	1.00	Gender, 1 if male (d)
Black	0.07	0.26	0.07	0.26	Race – black
Indigenous	0.01	0.09	0.01	0.09	Belongs to indigenous group
Education	3.64	3.23	0.00	17.00	Individual's years of education
Edu16	0.59	0.49	0.00	1.00	1 to 6 years of education (d)
Edu79	0.10	0.30	0.00	1.00	7 to 9 years of education (d)
Edu10	0.07	0.26	0.00	1.00	10 or more years of education (d)
Non-migrant	0.63	0.48	0.00	1.00	Has always lived in municipality (d)
Labor	0.40	0.49	0.00	1.00	Paid employee
Self-emp.	0.32	0.47	0.00	1.00	Self-employed
Unpaid	0.27	0.45	0.00	1.00	Unpaid worker
Hh size	5.15	2.52	1.00	35.00	Number of members in the household
Hh adults	3.45	1.71	1.00	23.00	Number of adult household members
Hh edu	2.34	1.77	0	12.8	Average years of education in household
Hh elec	0.71	0.45	0.00	1.00	Access to electricity (d)
Hh tel	0.06	0.24	0.00	1.00	Access to telephone line (d)
North	0.09	0.29	0	1	Residence in North (d)
North-East	0.41	0.49	0	1	Residence in North-East (d)
South	0.24	0.43	0	1	Residence in South (d)
South-East	0.21	0.41	0	1	Residence in South-East (d)
Center-West	0.05	0.22	0	1	Residence in Center-West (d)
Urban ext.	0.02	0.15	0.00	1.00	Residence in urban extension
Rural town	0.08	0.27	0.00	1.00	Residence in rural town
Municipal-level mean					
Mun_urban	0.60	0.23	0.00	1.00	Share of urban households in municipality
Mun_elec	0.74	0.27	0.00	1.00	Share of rural households with electricity
Mun_tel	0.06	0.09	0.00	0.93	Share of rural households with telephone
Mun_pop	30,847	186,763	795	10,400,000	Population in municipality
Pop1d	241,532	115,227	56,105	1,676,241	Distance weighted population, 1
Pop100d	629,104	1,398,288	246	16,200,000	Distance weighted population, 2
Dist500	258	195	2	1579	Distance to municipality with >500,000 people, km
Dist250500	206	174	2	1111	Distance to mun, 250-500,000 people
Dist100250	123	129	1	1275	Distance to mun, 100-250,000 people
Dist50100	76	74	1	900	Distance to mun, 50-100,000 people

Note: Variables indicated by (d) are dichotomous variables, taking value 1 if true, 0 otherwise. Continuous variables are converted into log-form in the model estimations. Sample size is 1,757,144. Municipal-level mean refers to unweighted mean for the 5,507 municipalities defined in the Demographic Census, 2000.

Thus the distance to the nearest population center with 500 000 or more people, for the average household in municipality *k*, was estimated by:

$$Dist500_k = d_k + \left(D_l \middle| Mun\_pop_l \geq 500\,000\right) \qquad (5)$$

given that municipality *l* is the nearest municipality with a population of 500 000 or more. The same method was used to estimate the other distance measures.

Estimation results

The results from the first of the three probability models are provided in Table 7. Explanatory variables were added stepwise to determine their explanatory power and their effect on other coefficient estimates. Human capital affects positively the probability of engagement in RNAE: Age has a positive and decreasing effect on the probability of non-agricultural employment, and the probability increases with level of educational attainment. As long as we do not differentiate between low- and high-productivity RNAE, women have higher probability of engaging in RNAE. Race matters little, even though people of indigenous ethnicity have a slightly higher probability of being involved in RNAE. Being a 'non-migrant' correlates negatively with RNAE, indicating that people who have moved are more likely to engage in non-agricultural activities.

When controlling only for personal characteristics, as in specification (*i*) in Table 7, having 10 years or more of education (as opposed to no education) increases the probability of RNAE by more than 50 percent. The marginal effects of all three educational levels decrease when household characteristics are added to the model, as indicated by column (*ii*). This gives support to the intra-household 'knowledge spillover' hypothesis: Given the individual's educational attainment, the education of other household members also influences employment outcomes. Household size has a small and negative effect on RNAE, which goes against the diversification-opportunity hypothesis; that is, even though larger households have more labor hours to allocate to various activities, this does not significantly affect the probability that a household member engages in RNAE as a principal occupation. To the extent that household access to electricity is an indicator of local infrastructural development, the hypothesis of a positive effect on RNAE is supported.[5]

[5] As a robustness check, the model was estimated with municipal-level fixed effects. This increased the explanatory power of the model, but it did not affect to a large extent the coefficients of the individual and the household characteristics.

Specifications (*iii*) to (*vi*) in Table 7 include municipal-level variables along with two indicator variables for rural sub-category. The results in column (*iii*) gives a clear message: Although the effects of individual and household characteristics remain stable, local conditions matter. Living in a rural area that is an urban extension increases the individual's probability of RNAE by more than 50 percent, as opposed to living in the rural exclusive category; residence in a rural town increases the probability by more than 20 percent. Given the location of the individual's residence, the degree of urbanization of the municipality at large also matters: the higher share of urban households in the municipality, the higher the probability of RNAE. Similarly, the more developed the rural infrastructure (indicated by, *Mun_tel*, the share of rural households with a telephone line), the higher the probability of RNAE.

Column (*iv*) adds the effect of market size. Local market size is estimated by distance-weighted size of the local population (*Pop1d*). The marginal effect reported in the table tells us that an increase in population by 10 percent, for the average municipality, increases RNAE probability by 3.3 percentage points for the rural resident. The use of the alternative proxy for regional market size – distance-weighted population within 100 kilometers – gave similar results in terms of qualitative effects and statistical significance, and did not alter the marginal effects of other variables noticeably. It is evident that distance to large population centers matters. Column (*v*) shows that the greater the distance to the nearest urban center with 500,000 or more inhabitants, the lower the probability of engaging in RNAE, as indicated by the negative coefficient of *Dist500*. Column (*vi*) shows that distance also to smaller population centers matters, controlling for the distance to larger population centers.

The second approach takes the heterogeneity of the RNA sector into account by distinguishing between low- and high-productivity RNAE. Table 8 contains the results from the estimation of the two binomial probit models (2) and (2'), estimating separately the probability of engaging in the two categories of RNAE. The previous analysis showed that women had a higher probability of engaging in RNAE. When separating high- and low-productivity jobs, however, women are at a disadvantage in the selection process into high-productivity RNAE; the marginal effect on being male is positive for high-productivity RNAE and negative for low-productivity RNAE. The pattern is similar for black people, even though the marginal effect is low. Education at all levels tends to affect positively the selection process into both forms of RNAE, but the effect of education is stronger for high-productive RNAE. In fact, the effect of high education on the probability of engaging in low-paid RNAE is even negative in the more complete model specifications (*iii*) and (*iv*).

TABLE 7: Results of probit model (1) estimation: probability of 'any' RNAE.

	Marginal effects on probability					
	(i)	*(ii)*	*(iii)*	*(iv)*	*(v)*	*(vi)*
Age	0.013***	0.011***	0.010***	0.010***	0.010***	0.010***
	(77.15)	(65.29)	(60.17)	(60.02)	(58.87)	(59.26)
Age2	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(86.17)	(77.57)	(72.08)	(71.77)	(71.07)	(71.47)
Male	-0.139***	-0.143***	-0.148***	-0.149***	-0.149***	-0.148***
	(155.74)	(158.04)	(158.75)	(158.74)	(159.07)	(158.26)
Black	0.002	0.019***	0.007***	0.007***	0.003*	0.004**
	(1.55)	(11.77)	(4.39)	(3.92)	(1.66)	(2.44)
Indigenous	0.018***	0.052***	0.039***	0.089***	0.075***	0.088***
	(3.64)	(9.99)	(7.46)	(16.24)	(13.80)	(16.18)
Edu16	0.105***	0.081***	0.081***	0.079***	0.080***	0.079***
	(95.37)	(68.91)	(67.43)	(65.41)	(65.95)	(65.53)
Edu79	0.322***	0.258***	0.239***	0.240***	0.241***	0.241***
	(181.20)	(136.41)	(123.45)	(122.53)	(123.31)	(123.39)
Edu10	0.513***	0.437***	0.424***	0.426***	0.427***	0.428***
	(269.94)	(204.84)	(191.66)	(191.12)	(191.06)	(191.33)
Non-migrant	-0.061***	-0.050***	-0.014***	-0.014***	-0.022***	-0.024***
	(70.51)	(56.91)	(15.64)	(14.88)	(23.82)	(25.79)
North	-0.023***	0.062***	0.119***	0.410***	0.174***	0.191***
	(15.13)	(35.88)	(63.99)	(144.31)	(90.09)	(96.39)
North-East	-0.073***	-0.025***	0.026***	0.131***	0.016***	0.028***
	(70.87)	(23.20)	(21.05)	(89.51)	(12.63)	(22.59)
South	-0.129***	-0.137***	-0.112***	-0.013***	-0.074***	-0.072***
	(114.94)	(122.87)	(89.58)	(8.40)	(57.13)	(55.38)
Center-West	-0.091***	-0.070***	-0.032***	0.145***	-0.002	0.023***
	(50.46)	(37.29)	(16.25)	(56.74)	(1.10)	(10.34)
Hh size		-0.006***	-0.003***	-0.003***	-0.004***	-0.004***
		(32.16)	(16.62)	(15.45)	(19.71)	(19.70)
Hh edu		0.023***	0.016***	0.016***	0.016***	0.016***
		(85.48)	(59.19)	(58.37)	(59.31)	(59.65)
Hh elec		0.126***	0.075***	0.058***	0.058***	0.050***
		(125.31)	(69.45)	(52.48)	(52.81)	(44.46)
Urban ext.			0.562***	0.507***	0.502***	0.503***
			(132.82)	(106.27)	(110.06)	(110.61)
Rural town			0.214***	0.227***	0.217***	0.226***
			(133.57)	(138.82)	(133.04)	(137.76)
Mun_urban			0.157***	0.106***	0.115***	0.094***
			(75.55)	(50.20)	(54.82)	(44.10)
Mun_tel			0.569***	0.349***	0.333***	0.285***
			(99.83)	(60.27)	(57.79)	(49.15)
Pop1d				0.259***		
				(135.37)		
Dist500					-0.090***	-0.074***
					(159.22)	(106.97)
Dist250500						-0.036***
						(60.26)
Dist100250						-0.006***
						(9.18)
Dist50100						0.002***
						(3.17)
Pseudo-R^2	0.108	0.126	0.176	0.189	0.193	0.195
Observations	1,757,144	1,757,144	1,757,144	1,757,144	1,757,144	1,757,144

Note: The dependent variable is *RNAE* (dichotomous variable to indicate RNAE as opposed to agricultural employment). Marginal effects are reported with robust z statistics in parentheses and refer to the change in probability of engaging in RNAE, given a small change in a continuous variable or a discrete change in a dichotomous variable. Asterisks denote significance level: * significant at 10%; ** at 5%; *** at 1%.

When household characteristics were added to the analysis, household average education has a negative effect on the probability of engaging in low-productive RNAE, but a positive effect on engaging in high-productivity RNAE. The inclusion of municipal-level variables yields similar results as in Table 7: Local market size, degree of urbanization, and rural infrastructure affect RNAE positively, both low-return and high-return activities. Column (*iv*), shows that distance to large population centers has a similar effect on both high- and low-productivity RNAE.

The multinomial probit model (3) complements these two separate probit models by estimating simultaneously the probability of engaging in each employment category.[6] It avoids the problem of odd comparison groups as in the case of the $RNAE^{LOW}$ model (2). Results from three specifications are reported in Table 9: a base specification with personal and household characteristics and two specifications including the estimated population and distance effects. The results are largely consistent with those from the two binomial models presented in Table 8. The marginal effects reported in the table give the estimated change in the probability that the average rural worker engages in the respective employment category, given a small change in the explanatory variable (the change from 0 to 1 for the dichotomous variables).

Overall, the results suggest that human capital matters for the probability of having non-agricultural employment among the rural labor force, but that this is largely a matter of selection into high-productivity RNAE. Even though women have a higher probability of engaging in RNAE, the decomposition of RNAE into low- and high-productivity jobs shows that they are at a disadvantage in the selection process into high-productive RNAE. Clearly, there are also multiple 'demand side' factors that matter. The 14 percent of rural households that are situated in areas that are considered urban extensions or rural towns have a much higher probability of engaging in non-agricultural activities. Controlling for the location of this minority of households, the size of the local market has a positive effect on RNAE opportunities. The impression from Figures A1—A5 is largely confirmed in the results above. The shorter the distance to population centers, the more likely is a rural resident to be engaged in RNAE.

[6] Due to the computational complexity of the multinomial probit model, the sample size was reduced to a random sample of one percent (17,571 observations) of the full base sample. Even though this random sample provides summary statistics that are very similar to those of the full base sample (reported in Table 6), the standard errors in the estimation results are larger than they would have been using the full base sample.

TABLE 8: Results of probit model (2) estimation: probability of low- and high-productivity RNAE.

	i) base		*ii) household*		*iii) population*		*iv) distance*	
	Low	*High*	*Low*	*High*	*Low*	*High*	*Low*	*High*
Age	-0.004***	0.018***	-0.005***	0.016***	-0.005***	0.016***	-0.005***	0.016***
	(41.29)	(124.82)	(43.83)	(114.14)	(50.90)	(113.08)	(52.21)	(112.36)
Age2	0.000***	-0.000***	0.000***	-0.000***	0.000***	-0.000***	0.000***	-0.000***
	(18.36)	(114.46)	(20.94)	(106.86)	(28.49)	(105.36)	(29.56)	(104.89)
Male	-0.160***	0.034***	-0.160***	0.032***	-0.160***	0.034***	-0.160***	0.034***
	(227.92)	(54.82)	(225.30)	(52.36)	(226.70)	(56.06)	(227.61)	(55.62)
Black	0.018***	-0.016***	0.021***	-0.006***	0.015***	-0.012***	0.013***	-0.013***
	(14.92)	(13.81)	(17.75)	(4.67)	(12.29)	(9.94)	(11.01)	(11.19)
Indigenous	0.017***	-0.014***	0.025***	0.005	0.028***	0.013***	0.031***	0.010**
	(4.95)	(3.76)	(6.98)	(1.18)	(7.74)	(3.20)	(8.60)	(2.36)
Edu16	0.018***	0.084***	0.014***	0.068***	0.012***	0.067***	0.012***	0.066***
	(22.66)	(99.80)	(15.79)	(77.34)	(14.30)	(77.02)	(13.86)	(76.54)
Edu79	0.067***	0.288***	0.054***	0.221***	0.038***	0.207***	0.038***	0.206***
	(51.61)	(171.82)	(39.40)	(131.36)	(29.06)	(124.18)	(28.79)	(123.80)
Edu10	0.029***	0.536***	0.014***	0.435***	-0.004**	0.423***	-0.005***	0.420***
	(21.42)	(279.65)	(9.77)	(212.35)	(2.52)	(205.60)	(3.51)	(204.31)
Non-migr	-0.026***	-0.027***	-0.023***	-0.019***	-0.006***	-0.006***	-0.008***	-0.008***
	(41.26)	(43.30)	(35.94)	(30.75)	(9.75)	(10.20)	(13.37)	(12.95)
North	0.006***	-0.029***	0.034***	0.017***	0.138***	0.159***	0.090***	0.058***
	(5.75)	(27.15)	(27.75)	(14.00)	(67.34)	(73.25)	(64.80)	(41.72)
North-East	-0.059***	-0.003***	-0.047***	0.028***	0.002**	0.086***	-0.023***	0.041***
	(78.65)	(4.52)	(60.37)	(35.81)	(2.25)	(84.13)	(27.59)	(48.34)
South	-0.058***	-0.050***	-0.059***	-0.054***	-0.019***	-0.001	-0.025***	-0.024***
	(72.38)	(64.74)	(73.45)	(71.87)	(18.13)	(0.68)	(27.49)	(28.38)
C-West	0.007***	-0.079***	0.017***	-0.071***	0.086***	-0.011***	0.055***	-0.052***
	(5.11)	(67.05)	(12.31)	(58.72)	(48.43)	(6.89)	(37.37)	(39.84)
Hh size			0.000***	-0.006***	0.002***	-0.005***	0.002***	-0.005***
			(3.26)	(40.74)	(14.77)	(33.15)	(13.04)	(34.75)
Hh edu			0.001***	0.017***	-0.003***	0.014***	-0.003***	0.014***
			(3.35)	(94.85)	(15.44)	(80.91)	(16.18)	(80.31)
Hh elec			0.052***	0.066***	0.024***	0.045***	0.021***	0.043***
			(70.63)	(91.75)	(31.83)	(59.25)	(27.75)	(57.00)
Urban ext.					0.108***	0.070***	0.096***	0.085***
					(46.68)	(32.99)	(44.98)	(41.00)
Rural town					0.104***	0.083***	0.101***	0.079***
					(86.70)	(73.70)	(84.46)	(70.69)
Mun_urban					0.072***	0.023***	0.069***	0.029***
					(48.76)	(15.95)	(47.61)	(20.79)
Mun_tel					0.188***	0.049***	0.147***	0.056***
					(49.70)	(13.17)	(39.55)	(15.31)
Pop1d					0.051***	0.087***		
					(44.69)	(76.37)		
Dist500							-0.033***	-0.032***
							(86.51)	(87.14)
Pseudo-R^2	0.075	0.120	0.080	0.142	0.109	0.164	0.114	0.165
Obs.	1,757,144	1,757,144	1,757,144	1,757,144	1,757,144	1,757,144	1,757,144	1,757,144

Note: The dependent variables are $RNAE^{LOW}$ and $RNAE^{HIGH}$. Marginal effects are reported with robust z statistics in parentheses. For $RNAE^{LOW}$, the marginal effects refer to the change in probability of engaging in low-productivity RNAE (as opposed to agricultural work *or* high-productivity RNAE), given a small change in a continuous variable or a discrete change in a dichotomous variable. For $RNAE^{HIGH}$, the benchmark is employment in agriculture *or* low-productivity RNAE. Asterisks denote significance level: * significant at 10%; ** significant at 5%; *** significant at 1%.

TABLE 9: Results of Multinomial probit model (3) estimation: probability of employment category

	i) base			*ii) population*		
	Agricultural empl.	*Low-prod. RNAE*	*High-prod. RNAE*	*Agricultural empl.*	*Low-prod. RNAE*	*High-prod. RNAE*
Age	-0.0122***	-0.00542***	0.0176***	-0.0118***	-0.00566***	0.0174***
	(0.0017)	(0.0012)	(0.0014)	(0.0017)	(0.0012)	(0.0014)
Age2	0.001***	0.001***	-0.001***	0.001***	0.001***	-0.001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Male	0.133***	-0.155***	0.0219***	0.137***	-0.158***	0.0205***
	(0.0093)	(0.0081)	(0.0061)	(0.0094)	(0.0082)	(0.0061)
Black	-0.001	-0.005	0.005	0.005	-0.007	0.002
	(0.016)	(0.012)	(0.012)	(0.016)	(0.012)	(0.012)
Indigenous	0.0085	-0.0268	0.0183	-0.0463	-0.00935	0.0557
	(0.050)	(0.032)	(0.045)	(0.058)	(0.036)	(0.055)
Edu16	-0.0661***	0.00337	0.0627***	-0.0665***	0.00323	0.0633***
	(0.011)	(0.0092)	(0.0086)	(0.012)	(0.0092)	(0.0087)
Edu79	-0.236***	0.0377**	0.198***	-0.230***	0.0347**	0.195***
	(0.020)	(0.015)	(0.020)	(0.021)	(0.015)	(0.020)
Edu10	-0.430***	-0.0249*	0.455***	-0.424***	-0.0301**	0.454***
	(0.021)	(0.014)	(0.024)	(0.021)	(0.013)	(0.024)
Non-migr	0.0272***	-0.0192***	-0.00795	0.0135	-0.0103	-0.00313
	(0.0088)	(0.0070)	(0.0064)	(0.0089)	(0.0070)	(0.0064)
North	-0.0634***	0.0270**	0.0363**	-0.432***	0.168***	0.265***
	(0.018)	(0.014)	(0.015)	(0.027)	(0.027)	(0.031)
North-East	0.00622	-0.0439***	0.0377***	-0.170***	0.0374***	0.132***
	(0.011)	(0.0087)	(0.0085)	(0.015)	(0.011)	(0.012)
South	0.0799***	-0.0443***	-0.0356***	-0.0127	-0.0136	0.0264**
	(0.011)	(0.0084)	(0.0079)	(0.015)	(0.011)	(0.012)
C-West	0.0303*	0.0340**	-0.0643***	-0.147***	0.131***	0.0158
	(0.018)	(0.016)	(0.010)	(0.026)	(0.024)	(0.019)
Hh size	0.00751***	0.000779	-0.00829***	0.00563***	0.00196	-0.00759***
	(0.0019)	(0.0015)	(0.0015)	(0.0019)	(0.0014)	(0.0015)
Hh edu	-0.0180***	0.00228	0.0158***	-0.0144***	-0.0000704	0.0145***
	(0.0026)	(0.0021)	(0.0018)	(0.0027)	(0.0021)	(0.0018)
Hh elec	-0.0902***	0.0366***	0.0536***	-0.0530***	0.0155*	0.0375***
	(0.0097)	(0.0078)	(0.0071)	(0.010)	(0.0083)	(0.0077)
Urban ext.	-0.563***	0.323***	0.241***	-0.388***	0.225***	0.164***
	(0.027)	(0.030)	(0.028)	(0.042)	(0.035)	(0.031)
Rural town	-0.195***	0.108***	0.0866***	-0.215***	0.119***	0.0962***
	(0.017)	(0.015)	(0.013)	(0.017)	(0.015)	(0.014)
Mun_urban				-0.140***	0.0925***	0.0472***
				(0.021)	(0.016)	(0.015)
Mun_tel				-0.435***	0.334***	0.100**
				(0.060)	(0.043)	(0.041)
Pop1d				-0.254***	0.107***	0.147***
				(0.018)	(0.013)	(0.012)
Dist500						
Obs.	17,571			17,571		
Wald chi^2	2,436			2,675		
Pseudo log-l	-78,703			-76,473		

Note: The dependent variable is employment category, *e*, where *e* is i) agricultural work, ii) $RNAE^{LOW}$, or iii) $RNAE^{HIGH}$. Marginal effects are reported with standard errors in parentheses. For each independent variable, the marginal effect refers to the change in probability of being in employment category *e*, given a small change in a continuous variable or a discrete change in a dichotomous variable.

TABLE 9 (*continued*)

	Agricultural empl.	*iii) distance Low-prod. RNAE*	*High-prod. RNAE*
Age	-0.0116***	-0.00565***	0.0172***
	(0.0017)	(0.0012)	(0.0014)
Age2	0.001***	0.001***	-0.001***
	(0.0000)	(0.0000)	(0.0000)
Male	0.139***	-0.158***	0.0193***
	(0.0094)	(0.0082)	(0.0061)
Black	0.005	-0.008	0.003
	(0.016)	(0.012)	(0.012)
Indigenous	-0.0243	-0.0103	0.0346
	(0.055)	(0.035)	(0.050)
Edu16	-0.0646***	0.00223	0.0624***
	(0.011)	(0.0091)	(0.0087)
Edu79	-0.227***	0.0327**	0.194***
	(0.021)	(0.015)	(0.020)
Edu10	-0.422***	-0.0298**	0.452***
	(0.022)	(0.013)	(0.024)
Non-migr	0.0208**	-0.0150**	-0.00575
	(0.0089)	(0.0069)	(0.0064)
North	-0.186***	0.0952***	0.0910***
	(0.021)	(0.017)	(0.018)
North-East	-0.0543***	-0.00727	0.0615***
	(0.012)	(0.0095)	(0.0094)
South	0.0514***	-0.0321***	-0.0193**
	(0.012)	(0.0094)	(0.0090)
C-West	-0.00567	0.0595***	-0.0539***
	(0.020)	(0.017)	(0.011)
Hh size	0.00650***	0.00143	-0.00793***
	(0.0019)	(0.0014)	(0.0015)
Hh edu	-0.0147***	0.000131	0.0145***
	(0.0027)	(0.0021)	(0.0018)
Hh elec	-0.0539***	0.0151*	0.0388***
	(0.010)	(0.0082)	(0.0076)
Urban ext.	-0.386***	0.196***	0.189***
	(0.041)	(0.033)	(0.031)
Rural town	-0.206***	0.114***	0.0926***
	(0.017)	(0.015)	(0.013)
Mun_urban	-0.156***	0.0931***	0.0630***
	(0.021)	(0.016)	(0.015)
Mun_tel	-0.423***	0.293***	0.130***
	(0.059)	(0.042)	(0.040)
Pop1d			
Dist500	0.0835***	-0.0439***	-0.0396***
	(0.0056)	(0.0043)	(0.0039)
Obs.	17,571		
Wald chi^2	2,747		
Pseudo log-l	-76,351		

IV. DETERMINANTS OF RURAL NON-AGRICULTURAL INCOME

We turn now to the analysis of determinants of income for the rural labor force. The same data were used for the income analysis as for the probit models. Income regressions were estimated separately for the agricultural and non-agricultural labor force to first determine what are the key determinants of earned income in rural areas, and second, to assess to what extent the structural coefficients differ between agricultural and non-agricultural workers.

Estimation method

Non-agricultural income is observed only for people who are engaged in RNAE, which gives us a censored sample. Of the N number of total observations, only the people (n_1) with RNAE report non-agricultural income; the people (n_2) in agricultural activities report zero non-agricultural income. The results from the probit model in the previous section suggest that there are individual characteristics and other factors that determine the selection process into RNAE, so that non-agricultural workers differ systematically from agricultural workers. Therefore, using only the n_1 sample in the regressions would provide biased and inefficient parameter estimates. To correct for the sample selection process, we used all N observations in a two-step maximum-likelihood estimation of the Heckman (1979) sample selection model. Firstly, we assumed that the selection into RNAE is determined by model (1) in the previous section. Secondly, controlling for the selection process into RNAE, we assumed that potential non-agricultural income for all N rural workers can be modeled as a log-linear function of individual, household, and regional characteristics:

$$y_{ijk}^{NA} = X_{ijk}\beta_1^{NA} + H_{jk}\beta_2^{NA} + M_k\beta_3^{NA} + \eta_{ijk} \quad (4)$$

in which y^{NA} is the log non-agricultural income of the individual, and X, Z, and M are vectors containing explanatory individual, household, and municipal characteristics, and η is the residual. Using model (1) again to control for non-selection into agricultural work, we model potential agricultural income, y^{AGR}, in the same fashion as non-agricultural income:

$$y_{ijk}^{AGR} = X_{ijk}\beta_1^{AGR} + H_{jk}\beta_2^{AGR} + M_k\beta_3^{AGR} + \eta_{ijk} \quad (4')$$

A third of the agricultural labor force is self-employed, and another third is, in the Census, classified as unpaid workers. This reflects the fact that the majority of people who work in agriculture are part of peasant households. A complete analysis of agricultural income determinants would have included a farm production function, which takes productive assets other than household labor into account. Given that our primary purpose of model (4') is to compare to what extent the effects of the explanatory variables differ for agricultural and non-agricultural income, and given the data limitations of the Census regarding productive assets, we applied the same specifications for the two income models.[7]

The explanatory variables were included on the same basis as in the probit analysis. In the probit model, we assumed that these variables determined the latent variable *v*, which could be thought of as the potential net gain of non-agricultural work. In this case, we estimated to what extend these variables determine actual earnings.

Results

Most of the factors that positively affect the selection process into rural non-agricultural employment also positively affect earned income. Table 10 provides the results from the estimation of model (4), with three specifications, which follow closely the setup of the probit analyses.

Controlling for other characteristics, men in the rural labor force have higher earnings than women, and the gender earnings gap is twice as large in the non-agricultural sector than in the agricultural sector. This is most likely a result of the selection mechanism discussed in the previous section: Women are more likely to engage in the low-paid forms of non-agricultural work. Blacks are at a disadvantage, with a negative marginal effect of about 8 percent on earnings. The human capital proxies – age and education – are of the expected sign: There are positive returns in both agricultural and non-agricultural work, at all three educational levels. The difference in returns to education is quite small between the sectors: Having ten or more years of education as opposed to none, has an effect on agricultural earnings of about 72 percent; for RNAE, the effect is a few percentage points higher. Household education also correlates positively with individual earnings, indicating that knowledge spillovers could have some effect not only in the employment outcome but also in the earnings potential. Table 4 show average earnings for self-employed are considerably higher than for laborers. This earnings differential is not well evident once we control for other characteristics. The effect is close to zero, and for people in non-agricultural activities the effect of being self-employed is negative, even though very small.

---

[7] In future research we plan to include a measure of household wealth as a proxy for the productive capital of self-employed agricultural households.

Earnings prospects are lower in the North-East than in the South-East (the benchmark region), and this effect is large (40 percent or higher) even when controlling for local characteristics. The results in columns (*ii*) and (*iii*) suggest that local characteristics affect not only employment outcome but also income prospects. Interestingly, the effect of local market size, degree of urbanization, and rural infrastructure are larger on agricultural than on non-agricultural earnings. This could be partly a result of locational variables mainly affecting the employment outcome, and once someone is engaged in the RNA sector, the returns to those activities are less dependent on location than are returns to agricultural activities. For agricultural activities, the size of the local market matters (*Pop1d*). Distance to population centers with over 500,000 people affects non-agricultural income negatively, and the positive effect on agricultural income is quite small and thus probably has little economic bearing. Specifications including the other distance measures (distance to municipalities with 50–100,000, 100–250,000, and 250–500,000) give a negative effect on earnings in agriculture, reinforcing the result that distance to local demand centers matters.

Table 10: Regression results of income model (4): agricultural and non-agricultural earned income

	i) base		*ii) population*		*iii) distance*	
	Agr.	*Non-agr.*	*Agr.*	*Non-agr.*	*Agr.*	*Non-agr.*
Age	0.027***	0.055***	0.028***	0.062***	0.028***	0.061***
	(57.72)	(93.36)	(63.75)	(105.25)	(62.86)	(103.89)
Age2	-0.000***	-0.001***	-0.000***	-0.001***	-0.000***	-0.001***
	(45.80)	(68.98)	(51.21)	(80.79)	(50.39)	(79.53)
Male	0.247***	0.630***	0.286***	0.569***	0.286***	0.575***
	(36.25)	(217.92)	(50.64)	(201.83)	(48.55)	(202.39)
Black	-0.092***	-0.089***	-0.080***	-0.084***	-0.082***	-0.085***
	(25.50)	(18.70)	(22.46)	(18.28)	(23.03)	(18.40)
Indigenous	-0.042***	-0.082***	-0.004	-0.043**	-0.036***	-0.054***
	(2.95)	(4.76)	(0.25)	(2.53)	(2.58)	(3.18)
Edu16	0.169***	0.183***	0.152***	0.226***	0.156***	0.221***
	(65.54)	(41.40)	(59.98)	(52.28)	(61.38)	(50.99)
Edu79	0.383***	0.332***	0.338***	0.428***	0.342***	0.416***
	(73.68)	(58.15)	(67.62)	(77.65)	(68.00)	(75.29)
Edu10	0.792***	0.645***	0.722***	0.800***	0.727***	0.780***
	(93.37)	(98.71)	(89.23)	(125.87)	(89.08)	(122.28)
Non-migrant	-0.111***	-0.105***	-0.105***	-0.102***	-0.100***	-0.100***
	(49.11)	(40.06)	(47.59)	(40.09)	(44.92)	(39.18)
Self-employed	-0.010***	-0.012***	0.018***	-0.012***	0.004*	-0.013***
	(4.87)	(3.72)	(8.15)	(3.64)	(1.87)	(4.06)
North	-0.028***	-0.029***	0.273***	0.118***	0.056***	-0.008
	(6.69)	(5.37)	(43.21)	(17.72)	(12.87)	(1.45)
North-East	-0.567***	-0.391***	-0.395***	-0.276***	-0.475***	-0.340***
	(209.71)	(113.20)	(121.03)	(72.09)	(168.42)	(99.62)
South	0.019***	0.049***	0.030***	0.079***	-0.043***	0.018***
	(6.00)	(14.76)	(8.61)	(19.87)	(13.08)	(5.29)
Center-West	0.172***	0.041***	0.315***	0.134***	0.183***	0.045***
	(41.73)	(6.91)	(62.51)	(20.35)	(44.32)	(7.66)
Hh education	0.086***	0.062***	0.078***	0.064***	0.079***	0.063***
	(114.41)	(84.42)	(106.60)	(88.72)	(107.11)	(87.37)
Urban ext.	0.475***	-0.035***	0.255***	-0.066***	0.293***	-0.048***
	(23.92)	(5.64)	(13.64)	(11.06)	(15.41)	(8.12)
Rural town	0.033***	-0.131***	0.040***	-0.022***	0.031***	-0.033***
	(7.02)	(27.97)	(8.76)	(4.82)	(6.87)	(7.41)
Mun_urban			0.120***	0.046***	0.156***	0.072***
			(22.92)	(8.00)	(29.91)	(12.46)
Mun_tel			1.395***	0.740***	1.531***	0.853***
			(80.92)	(51.82)	(87.99)	(61.33)
Pop1d			0.222***	0.129***		
			(45.58)	(34.22)		
Dist500					0.006***	-0.020***
					(4.24)	(15.10)
Constant	4.401***	3.890***	1.389***	1.742***	4.113***	3.515***
	(242.44)	(269.08)	(22.32)	(34.32)	(222.98)	(211.79)
Lambda	-0.31	-0.32	-0.28	-0.08	-0.27	-0.11
Observations	1,720,474	1,720,474	1,720,474	1,720,474	1,720,474	1,720,474
Censored obs.	952,503	1,253,852	952,503	1,253,852	952,503	1,253,852
Wald chi-2	193,171	154,113	206,368	171,617	204,509	169,280

Note: The dependent variables are log of earned income from principal employment. Coefficients are reported with robust z statistics in parentheses. Asterisks denote significance level: * significant at 10%; ** significant at 5%; *** significant at 1%.

V. CONCLUSION

In an attempt to assess to what extent rural non-agricultural employment constitutes a potential pathway out of poverty for people in rural Brazil, we analyzed the determinants of participation in the RNA sector and the earnings potential of jobs in this sector.

Employment opportunities in the RNA sectors depend on both supply-side factors (individual and household-specific characteristics) and demand-side factors (the local economic geography). Keeping the local economic conditions constant, people with higher education have a higher probability of engaging in non-agricultural activities. Education is also the key determinant that separates people who are engaged in high-productivity RNAE from those who are engaged in low-productivity activities. On the other hand, keeping individual characteristics constant, the local economic conditions fundamentally influence the non-agricultural employment opportunities. Firstly, aggregate local demand (estimated by distance-weighted population size) matters: the larger the local population the higher the probability of RNAE. Secondly, market access and transaction costs matter: the more urbanized the municipality and the higher the level of rural infrastructural development, the more people are engaged in non-agricultural activities. Thirdly, and as a consequence of the above, distance to large population centers matters.

Even though earnings, on average, are higher in non-agricultural sectors than in agriculture, the earnings potential in these sectors, just as the employment opportunities, are contingent on both supply and demand-side factors. Education is the single most important individual earnings determinant, and without education the income prospects for the rural workers are unlikely to be any better in the RNA sector than in agriculture. The separation of RNAE into low- and high-productivity activities shows that approximately half of the non-agricultural labor force has jobs that offer no higher income than the local agricultural wage rate.

In sum, RNAE is unlikely to be the appropriate pathway out of poverty for the majority of the rural poor, given that RNAE opportunities are lowest in locations where poverty is highest, and given that access to well-remunerated non-agricultural jobs depends on those personal assets that the poor are most likely to lack – human capital. In the larger context, however, it is evident that the rural non-agricultural sector is viable and important, given that there exist a certain level of aggregate local demand combined with a certain level of rural infrastructure.

REFERENCES

Barrett, C. B., T. Reardon, & P. Webb (2001), "Nonfarm Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics, and Policy Implications", *Food Policy*, 26(4), pp. 315–331.

Corral, L. & T. Reardon (2001), "Rural Nonfarm Incomes in Nicaragua", *World Development*, 29(3), pp. 427–442.

de Janvry, A. & E. Sadoulet (1993), "Rural Development in Latin America: Relinking Poverty Reduction to Growth", in Lipton & van der Gaag (eds), *Including the Poor.* World Bank, Washington D.C.

de Janvry, A. & E. Sadoulet (2001), "Income Strategies Among Rural Households in Mexico: The Role of Off-farm Activities", *World Development*, 29(3), pp. 467–480.

Dirven, M. (2004), "Rural Non-farm Employment and Rural Diversity in Latin America", *CEPAL Review*, vol. 83, August 2004.

Ellis, F. (2000) *Rural Livelihoods and Diversity in Developing Countries.* Oxford University Press, New York.

Escobal, J. (2001), "The Determinants of Nonfarm Income Diversification in Rural Peru", *World Development*, 29(3), pp. 497–508.

Ferreira, F. & P. Lanjouw (2001), "Rural Nonfarm Activities and Poverty in the Brazilian Northeast", *World Development*, 29(3), pp. 509–528.

Figueiredo, F., S. Helfand, and E. Levine (2007) "Income versus Consumption Measures of Rural Poverty and Inequality in Brazil", paper presented at *Pobreza Rural no Brasil: O Papel as Políticas Públicas*, conference in Brasilia, DF, April 17-18, 2007.

Heckman, J. J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1), pp. 153–161.

Helfand, S. & E. Levine (2004), "Changes in Brazilian Rural Poverty from 1991 to 2000: A Preliminary Analysis of Determinants, and an Assessment of Alternative Levels of Household Aggregation". Report prepared for OECD, 2004.

Isgut, A. E. (2004), "Non-farm Income and Employment in Rural Honduras: Assessing the Role of Locational Factors", *Journal of Development Studies*, 40(3), pp. 59–86.

Lanjouw, P. (1998), "Rural Nonagricultural Employment and Poverty in Ecuador", *Economic Development and Cultural Change*, 48(1), pp. 91–122.

Lanjouw, P. (2001), "Nonfarm Employment and Poverty in Rural El Salvador", *World Development*, 29(3), pp. 529–547.

Lanjouw, P. (2003), "Poverty and Non-farm Employment in Rural Brazil", in World Bank, *Rural Poverty Alleviation in Brazil – Toward an Integrated Strategy.* World Bank, Washington D.C.

Laszlo, S. (2005), "Self-employment Earnings and Returns to Education in Rural Peru", *Journal of Development Studies*, 41(7), pp. 1247–1287.

OECD (2005), *OECD Review of Agricultural Policies: Brazil*.

Verner, D. (2005), "Activities, Employment, and Wages in Rural and Semi-urban Mexico", World Bank Policy Research Working Paper 3561.

Verner, D. (2006), "Labor Markets and Income Generation in Rural Argentina", World Bank Policy Research Working Paper 4095.

World Bank (2003), *Rural Poverty Alleviation in Brazil – Toward an Integrated Strategy*. World Bank, Washington D.C.

World Bank (2006), "Brazil: Measuring Poverty Using Household Consumption", report number 36358.

Yang, D. T. & M. Y. An (2002), "Human Capital, Entrepreneurship, and Farm Household Earnings", *Journal of Development Studies,* 68(1), pp. 65–88.

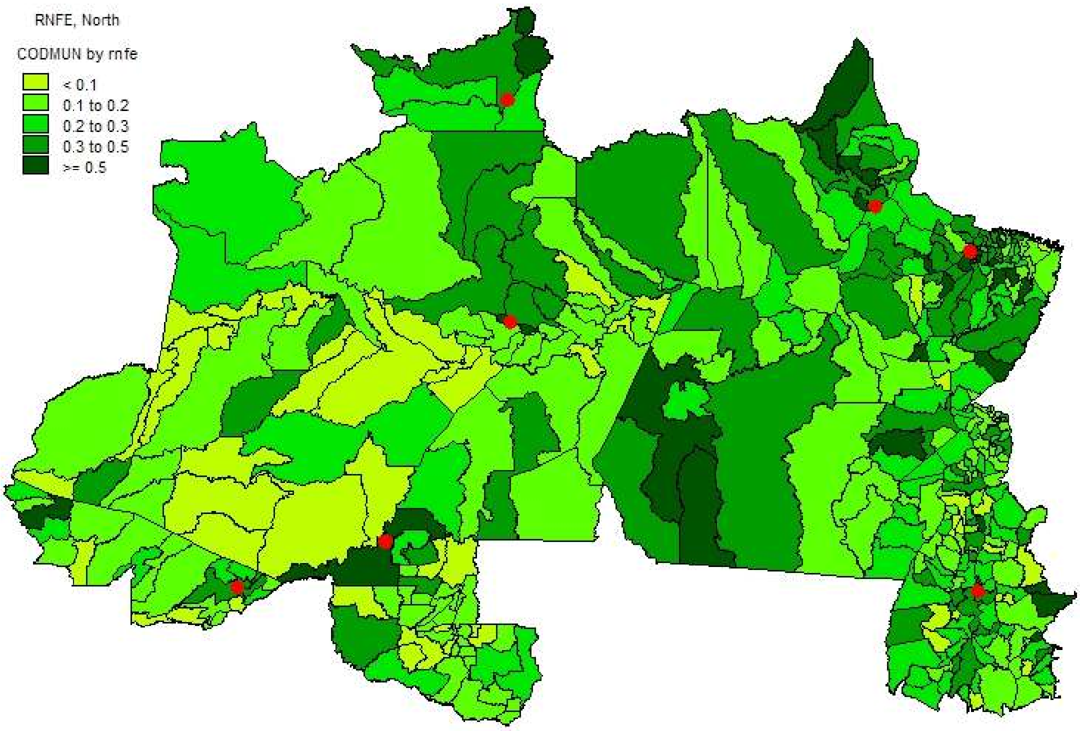
APPENDIX

TABLE A1: Demographic Census 2000: sample overview

	North	*North-East*	*South-East*	*South*	*Center-West*	*Brazil, total*
Full sample	1,522,222	5,966,953	8,267,281	3,121,695	1,396,261	20,274,412
Rural sample	518,204	2,201,473	1,112,372	782,023	243,738	4,857,810
Rural, adults	294,177	1,348,207	750,198	552,547	163,788	3,108,917
Rural, adult labor force	160,484	714,919	422,768	369,049	89,924	**1,757,144**

Note: Rural adult labor force sample is used in regressions. Adults are defined as people 15 years or older. The labor force consists of anyone reporting having a job (paid or unpaid).

Figure A1: Rural non-agricultural employment in Brazilian North



Source: Demographic Census 2000, authors' calculation and processing.
Note: Dots represent location of state capital city.

Figure A2: Rural non-agricultural employment in Brazilian North-East

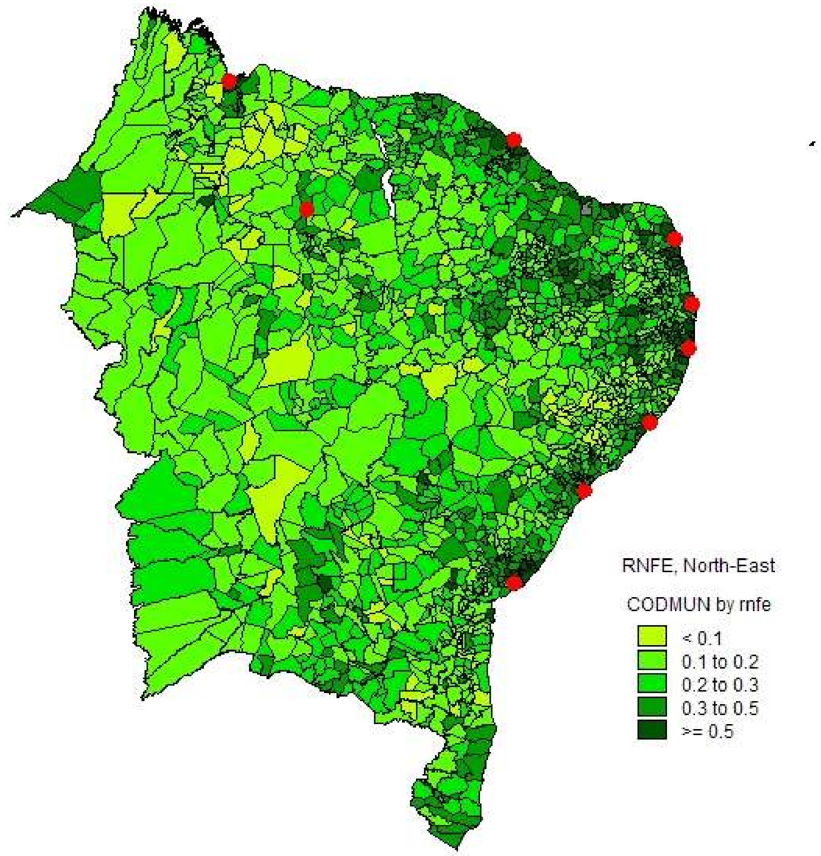


Figure A3: Rural non-agricultural employment in Brazilian South-East

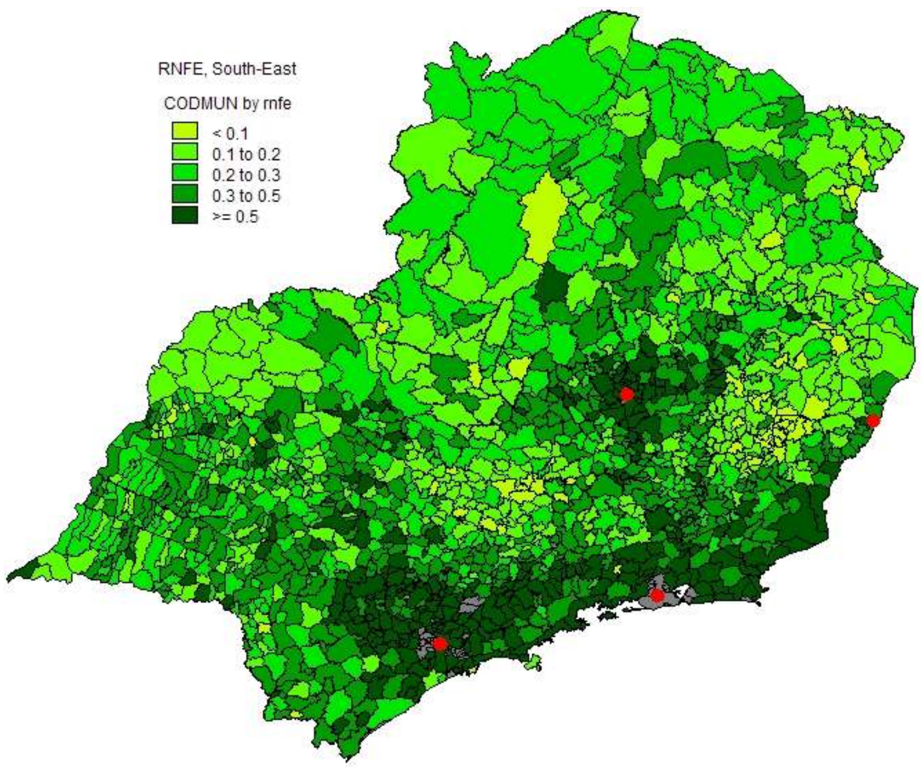


Figure A4: Rural non-agricultural employment in Brazilian South

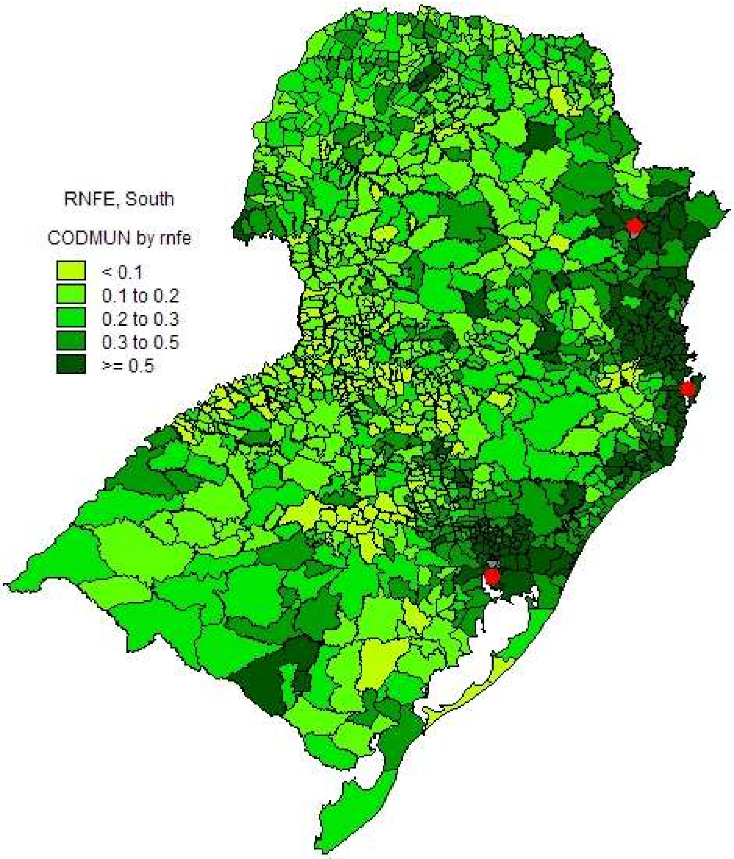


Figure A5: Rural non-agricultural employment in Brazilian Center-West