Agricultural Productivity and Family Farms in Brazil: Creating Opportunities and Closing Gaps

Steven M. Helfand,

Ajax R. B. Moreira, and

Edward W. Bresnyan, Jr.

Paper prepared for the World Bank
Sustainable Development Department
Brazil Country Management Unit
Latin America and the Caribbean Region

June 15, 2015

Table of Contents

List of Charts

List of Boxes

List of Figures

List of Tables

Executive Summary

Global population growth, rising effective demand for food, rivalry for water between its productive and consumption uses, and the imperative of limiting the expansion of the agricultural frontier in fragile ecosystems, all point to the urgency of raising agricultural productivity. Some 80-90 percent of the expected growth in agricultural production is expected to come from productivity gains in the form of increased yields and linked to technology adoption. Simultaneously, this essential productivity growth can reduce poverty, particularly in rural areas since agriculture is a source of income for more than 80 percent of the world's rural population. Smallholders or "family farms" constitute the vast majority of farm establishments in developing countries, with over 80 percent of these operating on less than two hectares.

Of the nearly 5.2 million farm establishments in Brazil, legally-defined family farms comprise 84 percent and account for 24 percent of the area of all farms and 38 percent of the value of agricultural production.[1] The remaining 16 percent of these establishments represent non-family agriculture. Fully one-half of Brazil's family farms are found in the Northeast macroregion; another 19 percent are found in the South, 16 percent the Southeast, 10 percent in the North and the remaining 5 percent in the Center-West. Brazilian agriculture overall, through strong capital investment and technological innovation, particularly in regard to agricultural research, has achieved a nearly 70 percent increase in value-added of crop production over the past decade. Export growth in beef, poultry, sugar and ethanol and soybeans testifies to this. Yet technological innovation has not been uniform across Brazil's macroregions, farm sizes or farm types.

Yield comparisons for specific products, taken from the 2006 Agriculture Census, indicate a consistent advantage of non-family farms over family farms in terms of land productivity. Forty-nine percent of Brazilian farms have 0-10 hectares; 92 percent of these are family farms. Capitalization varies widely on these farms: for the Northeast macroregion, non-land

[1] Under Brazilian Federal Law 11.326 of 2006, family farms have four defining characteristics. First, their land holdings are not more than four fiscal units (*módulos fiscais*), which vary in size across municipalities according to soil conditions and existing production patterns. In the Northeast macroregion, a fiscal unit ranges from 15 to 90 hectares, while in the South it varies between 5 to 40 hectares. Second, family farms derive a majority of household income from agriculture. Third, family farms primarily use household labor for on-farm activities. Fourth and finally, family farms manage on-farm activities themselves, i.e., no on-farm manager.

capitalization ranges from R$4,800 to R$7,500, while in the South and Southeast, non-land capitalization approaches R$50,000. In fact, the World Bank concluded in 2003 that Brazilian small farms and rural poverty was largely a question of, "those that rely on modern production technology and those that do not."

Using data from the 2006 Brazilian Agricultural Census, this Report assesses differences in agricultural productivity: (i) between two types of farms, namely family and non-family farms; and (ii) among four different sizes of these farms, as measured in hectares. Geography, technology, institutions and idiosyncratic characteristics all play a role in determining agricultural productivity. The Report pays particular attention to the role of public policies such as those that influence the use of credit, technical assistance, and irrigation, all of which are key to the technological change that is required to raise family farm agricultural productivity.

The Report focuses the analysis on the Northeast and South macroregions. Initially, a measure of poverty is constructed and applied to: (i) income solely from agricultural production; (ii) income from production and other earned income e.g., off-farm labor; and (iii) income from these two sources plus transfers e.g., *Bolsa Familia*. Chart A shows these three measures across farm types and sizes in these two macroregions.

Chart A: Poverty (%) by Size, Type, and Income Source

	Northeast Farm Size (ha)				**South** Farm Size (ha)			
Type/Income sources	0-5	5-20	20-100	100-500	0-5	5-20	20-100	100-500
Family:								
Production	81	67	60	54	62	35	30	-
Prod. + labor	79	65	58	52	57	33	28	-
Prod. + labor + transfers	65	51	44	41	43	26	23	-
Non-family:								
Production	81	62	51	51	58	37	38	51
Prod. + labor	53	46	42	47	26	23	33	47
Prod. + labor + transfers	44	39	37	43	22	21	30	44

Source: Agricultural Census-IBGE (2006)

Between the two regions and based on agricultural income alone, family farms in the South with zero to five hectares of land had 19 percentage points less poverty than farms of the same size class in the Northeast. The most important difference vis-à-vis poverty between family and non-family farms of this size was earning off-farm income. Nationwide, 45 percent of non-family farmers had off-farm earnings, whereas only 26 percent of family farmers did. When all income sources were considered, family farms with zero to five hectares in the South had 22 percentage points less poverty than those in the Northeast, the difference of which is almost entirely due to differences in agricultural income. An analysis of farm size, poverty and land productivity in the Northeast offers the following observations:

- *Land distribution is a key obstacle*: Yield increases alone for farms in the Northeast that have zero to five hectares (nearly one-half of all farms in this macroregion) will not be sufficient to lift them out of poverty;
- *Both larger farm sizes and higher levels of productivity* must combine to bring down poverty in the Northeast; and
- *Farm income alone is unlikely to resolve poverty* for agricultural producers in this region. Multiple policy approaches are necessary, including land reform, technical change, enhancing off-farm income opportunities, education, and transfers.

In the South macroregion, for farms with zero to five hectares, poverty falls from 77 percent to 47 percent with successive increases in land productivity. While agriculture is also not the only pathway out of poverty, plays a more important role in the South than in the Northeast.

Chart B: Value of Output per Hectare (R$/ha) by Farm Type and Farm Size

Type / Size(ha)	Mean Yield (R$/ha)		Ratio: F/NF
	Family (F)	**Non-Family (NF)**	
	648	**354**	**1.83**
0-1	7,192	13,263	0.54
1-5	2,858	7,315	0.39
5-10	1,867	4,224	0.44
10-20	1,198	2,794	0.43
20-50	668	1,877	0.36
50-100	308	1,019	0.30
100-200	119	685	0.17
200-500	82	352	0.23

Source: Agricultural Census-IBGE (2006)

Low productivity is related to insufficient levels of physical capital, purchased inputs, and human capital. This Report refutes the stylized notion that family farms have higher land productivity than non-family farms. In every farm size category, data indicate that non-family farms generate more value of output per hectare than family farms (Chart B). This result holds for the country as whole and within specific macroregions, such as the South and Northeast. Furthermore, every farm size, non-family farms use more purchased inputs and capital per hectare, which more than compensates for their use of less family labor. These differences in input intensities help to explain the higher yields for non-family farms at each farm size, and suggest that many family farms operate in a "low-level equilibrium".

A typology of family and non-family farmers also emerges from the 2006 Census data. Most family farmers are men (86 percent of total) between 35 and 65 years old with little formal education and more than ten years of experience running the current farm. Female family farmers (14 percent of total) are 7 percentage points more likely to be illiterate than their male counterparts, but similar to men in that most have primary education incomplete or less. Female family farmers are some 8 percentage points more likely to be at least 55 years old than are male family farmers. For family farmers, higher levels of schooling are also correlated with increased use of credit and technical assistance, fertilizers, irrigation, and specialization in production. Relative to family farmers, non-family farmers are quite different. They are 6 percentage points less likely to be over 55 years of age, 7 percentage points less likely to have at least 10 years of experience on the current farm, and 23 percentage points more likely to have at least a complete primary education.

The productivity of Brazil's family farms is an important determinant of their income, and of poverty in rural areas. This Report estimates two sets of models in order to analyze the determinants of productivity. First, a base model is estimated for all farms in Brazil. It assumes that the relevant dimensions of technological heterogeneity are: 1) farm type (i.e., family versus non-family); and 2) farm size (i.e., four farm size classes in terms of hectares). The model allows the technology to vary across the eight possible combinations of farm type and farm size.

A second group of models is estimated for family farms only. In this case, a translog production function is also estimated, but with technological heterogeneity by farm size and family farm "type", where type is alternatively defined by: (i) use of credit; (ii) technical assistance; (iii) use of irrigation; or (iv) specialization in production. Drawing on this second model, we also conduct counterfactual simulations to explore what would happen to family farms' productivity under a variety of scenarios.

First, we decompose the differences in productivity between the mean of each farm type and size of farm establishment into components attributable to differences in: (i) technology; (ii) non-factor characteristics that influence outcomes; and (iii) intensity of input usage. The non-factor characteristics include environmental factors, distance to markets, education, experience, and access to public goods and policies. The sum of the technology and non-factor components can be interpreted as the difference in total factor productivity (TFP) between any two groups. Second, we conduct a similar exercise for all "representative farms". For each of these farms, we simulate the productivity gain through use of a different technology. The simulations can be conducted unconditionally, or conditional on farm size or type. The results include average productivity gains for each group, and the distribution of the choices that optimize these gains.

Due to issues associated with the confidentiality of the microdata, all of the data used were aggregated by municipality, farm size in hectares (i.e., 0-5, 5-20, 20-100 and 100-500), and class (i.e., family, non-family). Data aggregation implies that we assume homogeneity within each aggregate observation, for example family farms with five to ten hectares in the municipality of Viçosa, MG. These are "representative farms." The econometric analysis explores variance between representative farms, but due to the aggregation, cannot examine variance within them.

The Report estimates these models in two macroregions of Brazil: (i) the Northeast, specifically in the Semi-Arid; and (ii) the South. The observations in the dataset total 7,144 representative farms in the Semi-Arid Northeast and 6,821 in the South, which aggregate the information from 1.481 million farms in the Semi-Arid and 926,000 farms in the South, which together account for nearly 50 percent of all farm establishments nationwide (Chart C). On average, there are around 45 individual farms in each representative non-family farm in the Semi-Arid, and around 38 individual farms in the South. Representative family farms have around 350 and 240 farms contained in each observation in the Semi-Arid and South, respectively, although there are far fewer in the 100-500 hectare class.

Chart C: Representative Farms by Type and Size (Semi-Arid Northeast and South)

	Semi-Arid Northeast			South		
	Representative farms	Farms (1000s)	Farms / Representative farms	Representative farms	Farms (1000s)	Farms / Representative farms
Total	7,144	1,481	207	6,821	926	136
Family (F)	3,796	1,330	350	3,282	790	241
0-5	964	705	732	1,054	190	180
5-20	988	346	350	1,073	380	354
20-100	1,027	247	240	1,104	218	198
100-500	817	32	39	51	1.9	36
Non-Family (NF)	3,348	151	45	3,539	136	38
0-5	671	54	81	675	19	28
5-20	814	32	39	924	29	31
20-100	922	34	37	1,023	39	38
100-500	941	31	33	917	49	54

Source: Agricultural Census-IBGE (2006)

On average, family farms have higher land productivity in both the Semi-Arid Northeast and the South; in the Semi-Arid region, land productivity of non-family farms is about 80 percent of the level of family farms, and in the South it is only about one-half. However, when we control for farm size, non-family farms use much more purchased inputs – including hired labor – and capital per hectare than family farms. In the Semi-Arid Northeast region, the differences are never less than 45 percent, and are often above 100 percent. Non-family farms in the South used 25 percent to 30 percent more purchased inputs per hectare than family farms in each size class, and they used between 20 percent and 66 percent more capital per hectare.

In both regions, non-family farms with zero to five hectares of land used more purchased inputs and capital per hectare and similar amounts of labor, yet land productivity was almost identical between non-family and family farms, suggesting an efficiency advantage for family farms of this size. For the larger farms, between 5ha and 100ha, the more intensive use of purchased inputs and capital is sufficient to produce higher land productivity for non-family farms in spite of less family labor being used. This is especially true in the Semi-Arid.

In terms of profit per unit of family labor, Semi-Arid Northeast family farms with zero to five hectares generated R$700, equivalent to slightly less than one-half of the annual poverty line (R$1,475). Family farms in this region with five to 20 hectares fared somewhat better, generating R$1,143 per unit of family labor, yet still did not surpass the annual poverty line. In the South, profit for family farms with zero to five hectares was nearly double that of the same farm size in the Semi-Arid Northeast (R$1,287) or 87 percent of the annual poverty line. For Southern family farms with five to 20 hectares, profit exceeded R$4,000 per unit of family labor, principally due to more intense use of purchased inputs and greater non-land capitalization.

In both regions regarding output elasticities with respect to inputs, the elasticities for non-family farms are larger than for family farms of the same size. Furthermore, capital elasticities tend to rise with farm size. In the Semi-Arid Northeast, purchased input elasticities for family farms (i) almost always fall with farm size; and (ii) are less than those for capital at all farm sizes. In the South, in contrast, purchased input elasticities (i) rise with farm size; and (ii) are uniformly less than those for capital at all farm sizes. Finally, for non-family farms in both regions, both capital and purchased input elasticities exceed those for family farms across all farm sizes.

In both the Northeast and the South regions, the evidence indicates that non-family farms use inputs more intensively, and have advantages with non-factor inputs such as human capital or climate and soil, but family farms off-set these disadvantages with higher levels of Total Factor Productivity (TFP). Family farmers close some of the yield gap by more efficiently turning inputs into outputs. Nevertheless, with their limited land and low levels of capital and purchased inputs, many family farmers, especially in the Northeast, cannot generate sufficient income to escape poverty.

Next, the Report turns exclusively to an analysis of family farms in both the Semi-Arid Northeast and the South macroregions and by farm size, exploring their differences across the following dimensions: (i) use of credit; (ii) uptake of technical assistance; (iii) specialization of production; and (ii) use of irrigation.[2] For each classification, we present information on four fronts: (i) poverty; (ii) use of inputs; (iii) factors associated with the selection into each group; and (iv) productivity, measured in both R$ per hectare and R$ per unit of family labor.

[2] For irrigation, due to data anomalies, only a descriptive presentation is made.

Chart D: Poverty Rate and Non-Factor Variables (Semi-Arid Northeast and South)

	Semi-Arid Northeast		South	
	Poverty Rate (%)	% Total Farms	Poverty Rate (%)	% Total Farms
Family	73	90	40	85
Credit				
Yes	68	15	24	38
No	74	85	50	62
Technical Assistance				
Yes	60	8	26	48
No	75	92	53	52
Credit and TA				
Both	56	2	20	26
Neither	76	79	59	40
Specialized[2]				
Yes	66	54	36	56
No	75	46	35	44
Irrigation				
Yes	57	6	29	5
No	75	94	41	95

Source: Agricultural Census – IBGE (2006)

Use of credit is associated with lower poverty rates in both regions, yet this association is much more profound in the South vis-à-vis the Semi-Arid Northeast (Chart D). The finding is all the more striking, given that 15 percent of Semi-Arid family farms used credit, whereas 38 percent used credit in the South. Technical assistance, which was accessed by only 8 percent of Semi-Arid family farms, was associated with a poverty rate of 60 percent, or a reduction of 13 percentage points. In the South, 48 percent of family farms accessed technical assistance and were associated with a poverty rate of 26 percent. Specialization was also associated with lower poverty rates in both regions, but the association was stronger in the South. Irrigation, uniformly sparse in both regions, was associated with the strongest reduction in the poverty reduction in the Semi-Arid Northeast, beyond that of credit, technical assistance or specialization.

Chart E draws comparisons for family farms by disaggregating the effects of credit, technical assistance, and specialization with respect to output, purchased inputs, capital and profits. For the Semi-Arid Northeast, disaggregation is also performed with regard to irrigation. In the Semi-Arid Northeast, family farmers that use credit generate yields (i.e., land productivity as measured in R$ per hectare) that are between 30 percent and 43 percent higher as a result of greater use of capital, inputs, and family labor. On average, these variables are 23 percent, 9 percent and 16 percent higher, respectively, than for family farms that do not use credit. As a result, on average, short-run profit per family member is 34 percent higher for farms using credit.

In the South, land productivity is 60 percent higher for those farms that use credit, and most of the yield gap comes from greater use of purchased inputs. Input usage is 40 percent higher for farms using credit.

Chart E: Non-factor variables and productivity

	Output	Purchased Inputs	Capital	Family labor	Profit (R$ per unit of family labor)
	(R$ per hectare)				
Semi-Arid Northeast:					
Credit (C)	536	147	1,054	0.30	1,317
No Credit (NC)	384	135	855	0.25	982
Ratio C/NC	1.4	1.09	1.23	1.16	1.34
Technical Assistance (TA)	621	156	1,288	0.22	2,082
No Tech. Assist. (NTA)	382	138	837	0.27	919
Ratio TA/NTA	1.63	1.12	1.54	0.84	2.27
Specialized (SP)	564	126	955	0.25	1,754
Not Specialized (NSP)	336	125	930	0.31	671
Ratio SP/NSP	1.68	1.01	1.03	0.80	2.61
Irrigation (I)	1,111	209	1,890	0.28	3,246
No Irrigation (NI)	358	143	813	0.26	834
Ratio I/NI	3.11	1.46	2.33	1.08	3.89
South:					
Credit (C)	2,005	865	4,753	0.19	6,082
No Credit (NC)	1,251	618	4,189	0.21	2,981
Ratio C/NC	1.6	1.4	1.13	0.88	2.04
Technical Assistance (TA)	2,082	808	5,187	0.18	6,937
No Tech. Assist. (NTA)	986	570	3,526	0.22	1,856
Ratio TA/NTA	2.11	1.42	1.47	0.82	3.74
Specialized (SP)	2,060	663	4,712	0.21	6,731
Not Specialized (NSP)	1,316	648	4,440	0.21	3,226
Ratio SP/NSP	1.57	1	1.06	1.00	2.09

Source: Agricultural Census-IBGE (2006)

The Report next explores the marginal effects of several key variables on the probability of: (i) being a family farmer; (ii) choosing to use credit and technical assistance; or (iii) specializing in production. Because the marginal effects are estimated by including farm size and other variables, they control for differences in many dimensions and thus compare producers that are similar based on observables. The marginal effects can be thought of as conditional correlations; they should not be interpreted causally.

In both the Semi-Arid Northeast and the South macroregions, the probability of being a family farmer *decreases* successively with increased schooling, and *increases* successively with increases in age and years in charge of the current farm. Women in both regions are more likely to be family farmers, are less likely to use credit and technical assistance, and are less likely to specialize in production. As the age of producers increases in both regions, the probability of using credit decreases and the probability of specializing increases.

The role of credit, technical and specialization were all decomposed for their relationship to land productivity and TFP for family farms. In the Semi-Arid Northeast, land productivity of family farms that used credit and technical assistance was explained by more intense use of purchased inputs, with little remaining to be explained by TFP. The importance of technology in explaining the TFP advantage among users of technical assistance in the South suggests that technical assistance is more associated with technical change than credit. Finally, the land productivity advantage of specialized farms is largely attributable to differences in TFP and likely the result of choosing higher-value crops.

Findings and Policy Implications

One of the main conclusions of this Report is that family farms have a high rate of poverty in Brazil because many of them have insufficient land and because they produce with extremely low levels of productivity. In the Northeast, more than one-half of family farms have between zero and five hectares of land and 81 percent of them do not generate enough farm income to lift their full-time equivalent family labor above the poverty line. A much lower share of family farms are poor in the South. This reflects differences across regions in productivity and in the land distribution. In the South only 24 percent of family farms have zero to five hectares of land, and this group achieves higher levels of productivity when compared to the Northeast. Based on farm income alone, 62 percent of this group in the South is poor. For farms with a bit more land—between five and 20 hectares—poverty based solely on farm income falls to 35 percent in the South, but only to 67 percent in the Northeast. Poverty reduction among family farmers—especially in the Northeast—requires policies that address both insufficient land and low levels of productivity.

Low productivity is related to insufficient levels of physical capital, purchased inputs, and human capital. In the South and Southeast, small farms (both family and non-family) with only zero to ten hectares of land had around R$50,000 of on-farm assets, while in the Northeast, farms of this size in Alagoas and Pernambuco had only 10 to 15 percent of this level of capital. Within regions, there was also ample evidence of differences in the use of capital between family and non-family farms. In the Semi-Arid Northeast, non-family farms used more than twice the capital per hectare used by family farms. Similar differences were observed in the use of purchased inputs within each region.

As a group, non-family farmers were younger and more educated: 39 percent of family farmers in Brazil were at least 55 years old, and 85 percent had not even completed a primary education. Non-family farmers were 6 percentage points less likely to be over 55 and 23 percentage points more likely to have a complete primary education or higher. These differences in schooling and the use of capital and purchased inputs contributed to the divergence in productivity and income across farms.

It is commonly accepted in Brazil that family farms have higher land productivity than non-family farms. Small farms have higher land productivity than large farms, and family farms tend to be small. When one compares farms of the same size, non-family farms uniformly have higher land productivity, although in the zero to five hectare class—where the distinction between family and non-family is less clear-cut—the difference might not be very large. The non-family farm advantage with regard to land productivity appears to be a result of a more intensive use of purchased inputs and capital, and higher levels of human capital, not of an inherent superiority of one type over another.

A second striking finding relates to the importance of off-farm work for non-family farmers. The inclusion of off-farm income lowered poverty by 28 percentage points, from 81 percent to

53 percent, for this non-family farms in the Northeast. For family farms, the poverty rate only dropped by 2 percentage points when off-farm income was included. The same phenomenon was observed in the South, where the inclusion of off-farm income lowered poverty by 32 percentage points for non-family farms, but by only 5 percentage points for family farms. Thus, non-family farmers not only hire labor in (one of the criteria for distinguishing between family and non-family farms), but some of them also hire out a significant amount of their own labor. The non-family farmers appear to have more education and skills, and thus are more competitive in labor markets. We suspect that they use this advantage to generate cash, relax credit constraints, and permit operating their farm at a higher level of productivity than family farmers. Thus, off-farm employment might contribute to poverty reduction both through increasing income directly, and by permitting higher levels of capital and purchased inputs to boost agricultural income on their own farms.

The pathways out of poverty include: (i) agriculture—either through intensification of family farms or wage labor; (ii) non-agriculture—either through labor market earnings or self-employment; (iii) migration, for those households that choose to exit from the sector; and (iv) transfers for those households without the potential to generate sufficient earned income. Where producers have sufficient land, poverty reduction depends on increasing productivity and income. In both regions that were studied in this Report, farms that used credit, technical assistance, irrigation or that specialized in production, often generated two to three times the profit per family member of farms of the same size that did not do so. As a result, poverty was significantly lower for these farms. In both regions, technical assistance was more strongly associated with gains in land productivity and income than was credit. This suggests that while credit can relax constraints and permit increased use of purchased inputs, technical assistance is more closely related to technical change and income growth. This finding reinforces the importance of investments in technical assistance to spur needed technology adoption among family farms as a means of stimulating TFP growth. This is especially important in risky climactic environments such as the Semi-Arid portion of the Northeast.

For family farms with insufficient land to permit movement out of poverty (e.g., 0-5 hectares in Northeast Brazil), land reform can be one piece of the solution, especially in the Northeast, along with off-farm labor opportunities. Policies that support access to off-farm earnings can play an important role in poverty reduction. In both regions that were studied in this Report, off-farm earnings reduced poverty for small non-family farmers by more than off-farm earnings and transfers combined for family farms of the same size. Numerous studies have shown that access to off-farm employment increases with human capital.

With regard to non-agricultural employment, proximity to population centers is another important factor that increases the probability of finding employment. There are many possible growth engines that can create labor demand for the poor. In some regions, the existence of irrigated agriculture might create linkages to non-agricultural jobs in fruit and vegetable processing. In other locations it might be an abundance of animal production that creates employment in a slaughter house. The diversity of possibilities for growth and employment has led naturally to a focus on territorial development. If alternatives to migration are to be constructed, local territories must find their own dynamic sectors that have the potential to generate employment. It is not essential that these sectors be based on agriculture. What matters is that they create opportunities that are accessible to the poor.

Among the highest priorities for public policy in Brazil should be to improve the quantity and quality of education for young people who live in rural areas. This is perhaps the only policy that contributes positively to all pathways out of poverty. Education is associated with higher agricultural income as a result of its relationship with productive efficiency, technological adoption, and the ability to participate in input and output markets. Education is also associated with higher non-agricultural income because it increases the probability of finding employment and the earnings of individuals once employed. Education is an important factor that contributes to more successful migration histories. Finally, education is a key component throughout the world in cash transfer programs that seek to break the transmission of poverty from one generation to another. In spite of these benefits, education for rural households continues to lag behind urban areas in both quantity and quality. The international literature has shown convincingly that the social returns to investments in education—especially at the primary level—are quite high. In order to reap the full long-term benefits of *Bolsa Família*, and provide a more promising future for rural youth regardless of the pathway that they pursue, Brazil needs to improve the quality of schools for its rural population.

It is often observed that farmers are not young in Brazil. While it is true that 39 percent of family farmers were over 55 years old in 2006, it is also true that 34 percent were between 25 and 45. The younger cohorts have more schooling, and appear to be more willing to experiment with new technologies. In order for agriculture to provide a pathway out of poverty for this group, it is essential that policies assist them to produce with sufficient land and at much higher levels of productivity than the previous generation of farmers. Nearly 50 years after Schultz's 1964 seminal publication, it is time to finally achieve "efficient and not poor."

1. **Introduction**

1.1 Population Growth and Demand for Food: the Global Challenge of 2050

Agricultural productivity is pivotal to achieving the challenge of feeding the estimated nine billion people that will inhabit the Earth in 2050. When compared to 2005/2007 levels, overall food production will need to increase by nearly 70 percent to meet this future demand (FAO 2009, Global Harvest Initiative 2011). Some 80-90 percent of the expected growth in agricultural production is expected to come from increased yields; the remainder will require an expansion in arable land, almost exclusively in developing economies. Achieving these yields will be closely tied to technology adoption. For example, irrigated lands will need to expand by nearly 20 percent – again primarily in developing economies – and water efficiency gains must accompany such an expansion as increasingly scarce freshwater must serve both human and agricultural needs.

The needed productivity growth can be a significant driver of poverty reduction, given that agriculture is a source of income for more than 80 percent of the world's rural population, and food price spikes hit hardest on the urban poor. Cross-country evidence has shown that GDP growth in agriculture can be at least twice as effective in reducing poverty as growth from other sectors (World Bank 2007). Smallholders or "family farms" represent the vast majority of farm establishments in developing countries, with over 80 percent of these operating on less than two hectares. As such, these small farms must play a crucial role in driving productivity growth, which can be mutually beneficial in meeting global food demand and reducing rural poverty.

1.2 Agricultural Productivity in Brazil: Conquests and Contrasts

Brazilian agriculture exemplifies both conquests and contrasts. Of the 5.175 million farm establishments nationwide, legally-defined family farms comprise 84 percent and account for 24 percent of the area of all farms and 38 percent of the value of agricultural production. The remaining 16 percent of establishments represent non-family agriculture (Box 1). These are just a few of many stark contrasts that depict the bi-model nature of the Brazilian agriculture sector and, in their totality, aid in understanding the challenges faced by family agriculture in terms of productivity and competitiveness.

While land concentration is one element of the bimodal nature of Brazilian agriculture, technology concentration perhaps is a potentially better explanatory factor for the divergence in productivity and incomes between family and non-family farms. In the presence of production technologies and management practices that can accelerate yields on limited land holdings, it is access to such innovations that becomes crucial, particularly for the family farm.

Box 1: Family Farms in Brazil

Under Federal Law 11.326 of 2006, family farms have four defining characteristics. First, their land holdings are limited to four fiscal units (*módulos fiscais*) which vary widely in size across municipalities according to soil conditions and existing production patterns. In the Northeast macroregion, a fiscal unit ranges from 15 to 90 hectares, while in the South it varies between 5 to 40 hectares. Second, family farm derive a majority of household income from agriculture. Third, family farms primarily use household labor for on-farm activities. Fourth and finally, the household manages on-farm activities itself.

Family farms gain legal recognition through the issuance of a Declaration of Eligibility for the Federal Program for Family Agriculture - PRONAF (*Declaração de Aptidão ao PRONAF - DAP*), which makes them eligible for participation in PRONAF and other programs.

The legal definition of a family farm is used throughout this report due to its importance as a reference for public policy in Brazil. The classification does, however, have some shortcoming in terms of identifying a socioeconomic category. In many parts of the world, family farms are identified by type of management alone—those that are owner-operated—without reference to other factors such as work off-farm. This creates a discrepancy with the international literature—and even with some of the Brazilian literature—that must not be forgotten. For example, Brazilian small farms that are well-inserted into the labor market and earn more than half of their income off-farm, are identified legally as non-family farms. In many other parts of the world, these would be classified as diversified family farms. Even in Brazil, until recently many researchers would have described them as "pluri-active households." In spite of these limitations, throughout this report the legal definition is used due to its growing importance for public policy.

Source: IBGE 2009

Over the past several decades, Brazil has become a globally competitive powerhouse in agriculture. In the past decade alone, it has achieved a nearly 70 percent increase in the value-added of crop production and a ten-fold increase in beef exports. Brazil is now the world leader in exports of poultry, sugar and ethanol and second worldwide in soybean exports. In fact, agriculture accounts for about 36 percent of total exports, primarily from top commodities such as coffee, soybeans, beef, orange juice concentrate and sugarcane-derived ethanol. These impressive results reflect a strong history of capital investment and technological innovation focusing on commercial agriculture, particularly in regard to agricultural research. The transformation of the Brazilian *Cerrado* – straddling the Northeast and Center-West macroregions – is a testament to such innovation. Indeed, a World Bank 2003 analysis of small farms in the context of rural poverty nationwide concluded the following:

> "The division that comes to be relatively more applicable is not between the groups that possess [more] land and those that do not, but rather, between *those that rely on modern production technology and those that do not* (emphasis added)."

1.3 Poverty in Brazil and its Link to Agriculture

During the first decade of the 21st century, Brazil achieved sizable reductions in poverty nationwide. Between 2003 and 2009, poverty fell by 40 percent and extreme poverty by 52 percent. More than 22 million Brazilians emerged from poverty over this period and almost 13 million exited extreme poverty (World Bank 2011). Brazil stands out among major middle income countries in that it has been able to combine economic growth with reduced inequality. Economic growth explains about one half of this large reduction in poverty (Figure 1). The remaining 50 percent is associated with a decline in inequality, which can be decomposed into three main factors: (i) labor income growth (45 percent of the fall in inequality); (ii) increased government transfers e.g., *Bolsa Família* (45 percent); and (iii) reduced dependency rates (10 percent).

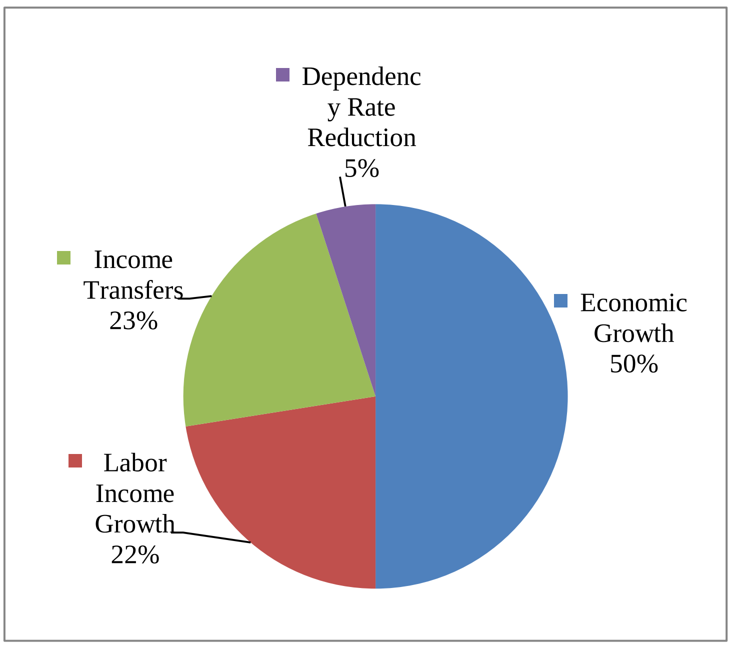


Figure 1: Sources of Poverty Reduction in Brazil

Source: World Bank (2011)

Despite these impressive gains, living standards for many millions of Brazilians remain far from OECD averages. Brazil is one of the 10 most unequal countries globally and many of the poor still lack access to quality public services, e.g., potable water, sanitation. More than 41 million Brazilians still live on less than US$3.50 per day and life expectancy at birth remains relatively low (72.3 years in Brazil, compared to 75.2 in Argentina and 78.5 in Chile). Regional inequalities also persist; the North and Northeast macroregions trail national average incomes by some 30 to 40 percent and exhibit much lower life expectancies (World Bank 2011).

Extreme poverty (i.e., with monthly per-capita income up to R$67 or US$35) still defines 8.6 million Brazilians, concentrated in rural areas and principally in small towns in the Northeast macroregion (Osorio et al., 2011). Some three million of these extremely poor (36 percent) make their living in agriculture. Much of this population receives government assistance; for example, 80 percent of them access *Bolsa Família*.[3] Small farm size for many family farms appears to be associated with such extreme poverty. In the Northeast macroregion, 20 percent of family farms have less than one hectare of land – while one-third have between one and five hectares and another one-third between five and 50 hectares. By contrast, in the South macroregion, 69 percent of family farms have between five and 50 hectares. According to Osorio et al. (2011), the major obstacles to exiting extreme poverty for those families connected to agriculture are: (1) small farm size; (2) access to inputs, including water and technical assistance, and (3) difficulty selling their products.

1.4 Closing gaps: Competitiveness, Innovation and the Family Farm

Productivity is a key element in gaining and maintaining competitiveness. In fact, Porter (2005) asserts "True competitiveness is measured by productivity." Yet productivity is but one element of a larger set of factors that determine competitiveness and, specific to this report, agricultural competitiveness. The World Economic Forum (WEF) puts forth twelve facets or "pillars" for determining competitiveness, progressing through three successive stages:

[3] A 2010 Impact Evaluation, conducted by IFPRI, found that in households receiving Bolsa Familia payments: (a) Infants were more likely to receive their vaccinations on schedule; (b) School attendance by boys and girls rises by 4.4 percentage points; for the Northeast macroregion, enrollments have risen by 11.7 percentage points; (c) Children (especially girls aged 15 to 17) are more likely to progress from one grade to the next; and (d) Pregnant women have 1.5 more pre-natal visits with a healthcare professional.

(i) *factor-driven* – where basic requirements are essential (e.g., sound infrastructure, functional institutions, stable macroeconomy, adequate health and education);

(ii) *efficiency driven* – requiring greater focus on higher education and training, strong financial intermediation and well-articulated output and input markets, including labor; and

(iii) *innovation-driven*, backed by an entrepreneurial class that can adapt and retool under dynamic market conditions.

The challenge ahead is to bring these theories to the data and assess their empirical relevance. As this Report will demonstrate in subsequent sections, commercial agriculture in Brazil has flourished in recent decades through progress across these three pillars, while family agriculture has yet to realize its full potential along these same dimensions.

Table 1: Yields (kg/ha) for selected crops by farm type

Crop	Family Farms (F)	Non-Fam Farms (NF)	NF/ F
Rice	2,741	5,030	1.84
Black bean	831	1,288	1.55
Manioc	5,770	7,541	1.31
Corn	3,029	4,303	1.42
Soybean	2,365	2,651	1.12
Wheat	1,480	1,822	1.23

Source: Agriculture Census (IBGE 2006)

Yield comparisons for specific products, taken from the 2006 Agriculture Census, indicate a consistent advantage of non-family farms over family farms in terms of land productivity (Table 1). Particularly for family farms, given their strong role in overall agricultural production, it would also be helpful to assess household labor productivity. Capital investment also appears to vary significantly between family farms and non-family farms and across regions. Nationwide, farms with 0-10 hectares — 92 percent of which are family farms — account for 49 percent of all farm establishments; capitalization for these farms (excluding land) averages R$15,523. Nonetheless, stark disparities exist between the Northeast macroregion – where most of the farms with 0-10 hectares reside – and the South and Southeast macroregions. For example, farms with 0-10 hectares in Alagoas and Pernambuco show non-land capitalization of R$4,888 and R$7,525, respectively, while in Santa Catarina and São Paulo, these small farms show non-land capitalization on the order of R$50,000.

Yet these partial measures of productivity for land, labor and capital are limited in their usefulness. It is well-established that small farms almost always have higher yields than large farms, but they usually have lower labor productivity. Higher land productivity does not necessarily mean higher total factor productivity, profits, or welfare. For this reason, a measure of Total Factor Productivity or TFP, which relates output to all of the inputs used in production, gives a superior indicator of a sector's efficiency than do indexes of partial productivity (Alston, Beddow, and Pardey 2010).

Alston, Beddow and Pardey (2010) report global indices of agricultural growth and TFP over the period 1961-2007 using aggregate data on grain outputs, inputs and yields from some 170 countries (Table 2). Their estimates indicate that global agricultural output grew at 2.8 percent annually in the 1960s and then maintained a fairly steady annual growth rate of slightly over 2 percent for each decade since 1970. Over time, an increasing share of output growth was due to improvements in TFP rather than input accumulation. Input growth slowed significantly, from over 2.3 percent per year in the 1960s to only 0.74 percent per year during 2000-07 (and even lower in the 1990s when agricultural severely contracted in the transition economies of the former Soviet Union and Eastern Europe). Improvements in TFP kept global output growth steady as the rate of input accumulation fell.

Table 2: Average Annual Global Agricultural Growth Rate by Period (%)

Period	Output	Input	TFP	Output per Worker	Output per Hectare	Grain Yield (t/ha)
1961-1969	2.81	2.31	0.49	0.96	2.39	2.84
1970-1979	2.23	1.60	0.63	1.46	2.21	2.62
1980-1989	2.13	1.21	0.92	0.97	1.72	2.00
1990-1999	2.01	0.47	1.54	1.15	1.74	1.61
2000-2007	2.08	0.74	1.34	1.72	2.10	1.01
1970-1989	2.18	1.40	0.77	1.22	1.97	2.31
1990-2007	2.04	0.59	1.45	1.40	1.90	1.35
1961-2007	2.23	1.24	0.99	1.25	2.01	2.02

Source: Alston, Babcock and Pardey (2010)

In the Brazilian context, TFP growth accounted for 72 percent of agricultural growth from 1975 to 2005 (Gasques et al. in Denegri and Kubota 2008). For the shorter period 2000 to 2005, it appears that TFP growth contributed 65 percent to overall output growth in the sector, while the increased use of inputs, especially land and capital, accounted for the remaining 35 percent of the increase in agricultural product.

In Brazil, annual agricultural TFP growth has averaged 3.77 percent from 1975-2011 (Gasques et al. 2012), comparatively higher than both that found by Ball (2006) for the United States and the global estimates presented in Table 2. But the period of higher productivity growth was from 2000 to 2011, when TFP grew by an astounding 5.69 percent per year. Three factors have spurred TFP growth and its associated agricultural output growth:

(i) changes in the composition of agricultural output, with products of animal origin and livestock increasing their share in value terms.

(ii) the marked expansion of rural credit in recent years, and its impact on access to new technologies and expansion of production scale. Undoubtedly, this aspect was reflected directly in the growth of both output and TFP. As we shall see below, rural credit expansion does not equate to equal access to rural credit across regions and types of producers.

(iii) Finally, agricultural research was crucial as a determinant of productivity gains in agriculture, with commercial agriculture as the primary beneficiary of such research (Denegri and Kubota 2008).

Understanding the determinants of technology adoption is key to explaining TFP variation whether across countries or within countries. There are a number of theories linking technology adoption to the role of institutions (Acemoglu et al., 2006), financial markets (Alfaro et al. 2006 and Aghion et al., 2006), endowments (Caselli and Coleman, 2006) and policies (Holmes and

Box 2: Innovation in Brazilian Agriculture as a Driver of Competitiveness

In its 2012-13 Report, the World Economic Forum's *Global Competitiveness Report* rated Brazil 39th among 144 countries when assessed in terms of innovation. The Report evaluates innovation along the following parameters: capacity to innovate, quality of scientific research institutions, corporate spending on research and development, university-industry research collaboration, government procurement of advanced technology products, availability of scientists and engineers, and utility patents per capita. The Report posits that in the current globalized world, competitiveness will determine who will be the real producers of knowledge, the suppliers of technologies and products, and the providers of services in the world market.

Innovation is often *location-specific*, as in the case of Brazil's conquest of the *cerrado* (Brazilian savannah) in the Central-West macroregion. Previously, there was no technology specific to agriculture in the *cerrado*, which required the adaptation of forms of agriculture used outside Brazil to this context and highlights the role of the National Agricultural Research Agency (Embrapa) in achieving this. Technological innovation in Brazilian agribusiness also benefited from policy instruments, among these agricultural credit policy targeting large-scale producers, that permitted the opening of new areas and the purchase of machinery and equipment to expand the amount of land being cultivated; also the guarantee of minimum prices reduced producers' risk.

In spite of the achievements, bottlenecks persist. Improving infrastructure, resolving sanitary standards, paying greater attention to land distribution and titling and investing in strong rural extension – particularly as these relate to smallholders – can broaden the benefits of innovation among farm establishments.

Source: Mia et al (2009); Global Economic Forum (2013)

Schmitz, 2001). As later chapters of this Report will demonstrate, these aforementioned facets play a role in productivity and, more specifically, remain to be fully exploited by a large segment of family farms in Brazil, particularly in Northeast Brazil.

Using the 1995-96 Agricultural Census, Helfand et al. (2011) report poverty rates that are 36 percentage points higher for family farms relative to non-family farms in Northeast Brazil, where 50 percent of the family farms are located. In the other macroregions nationwide, this gap ranged from 15 (South) to 26 percentage points (Southeast). There is also considerable heterogeneity within the family farm sector. Using a slightly different definition of a family farm, Guanziroli et al. (2010) classify 9.1 percent of the family farms in a high-income class ("type A"). This sub-group produced 68 percent of the value-added of family farm output in 2006, and earned an average net monetary income of R$53,000 per farm. The low-income "type D" family farms in their study represented 58 percent of all family farms, yet only produced 11 percent of output and earned only R$255 of net monetary income per farm (Table 3). Notably, while the average size of Type A family farms shrunk by one-third from 1996 to 2006, these farms saw their net annual incomes grow by 75 percent over the same period, which would seem to indicate a strong effect of intensification. Conversely, all other family farm types experienced negative income growth over the same period and all but the low-income type D farms showed a reduction in farm size. As such, by either definition, insufficient farm income is clearly an important problem for the vast majority of family farms.

Table 3: Family Farms: Value-added production, farm size and net income

Type of Family Farm	No. Family Farm Estab.	% Estab.	% Value-Added Production	Average Farm Size (HA)			Net Annual Income (R$)		
				1996	2006	% Change	1996	2006	% Change
A	412,806	9.1	68	59	39	-34	30,333	53,236	75
B	941,716	20.7	16	34	24	-29	5,537	3,725	-33
C	572,518	12.6	5	22	20	-9	1,820	1,499	-17
D	2,624,927	57.6	11	16	21	31	-265	255	-11
Total	**4,551,967**	**100.0**	**100**	**26**	**23**	**-11**	-	-	-

Source: Guanziroli et al. (2010)

1.5 Programs and Policies in support of Family Agriculture

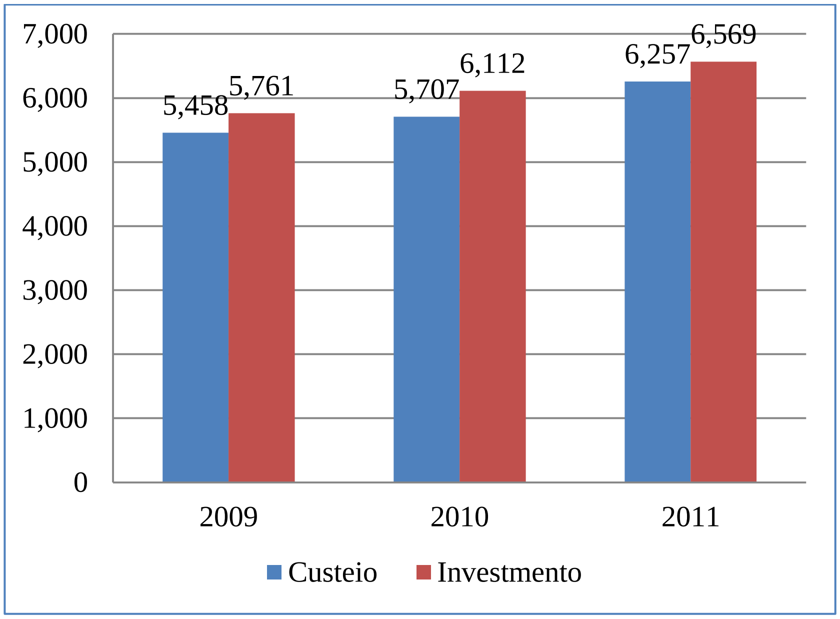
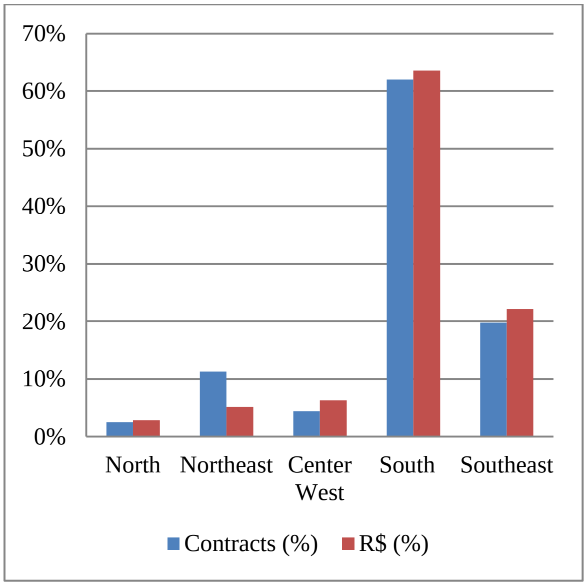


Figure 2: Rural Credit: Short- and Long-Term (2009-11, in R$million)

The Ministry of Agrarian Development (MDA) executes Brazil's Federal public policy for family agriculture, principally through its Secretariat for Family Agriculture (SFA). Assistance to family agriculture cuts across two main axes: (i) finance, the bulk of which is channeled through the National Program to Strengthen Family Agriculture – PRONAF; and (ii) the National Program of Technical Assistance and Rural Extension – ATER. Under PRONAF, family farms can access financing under two "windows": (i) for short-term, production-related expenditures (known as *custeio*); and (ii) for long-term investment in on-farm infrastructure, plant and equipment expenditures (known as *investimento*). From a total of R$2 billion in 2002, PRONAF has grown to a total program of R$12.8 billion in 2011 (Figure 2). In 2011, the value of *custeio* contracts totaled R$6.257 billion nationwide, with a mean value of R$9,582, while for investment contracts the figures were R$6.569 billion and R$7,632, respectively.

Yet access to PRONAF differs markedly, both across the two windows and among the five macroregions nationwide. In 2011, for both *custeio* and *investimento*, the South and Southeast macroregions dominated in terms of approved contracts and total value of loans. Together, these two macroregions captured 86 percent of *custeio* and 63 percent of *investimento*. By contrast, the Northeast macroregion, home to nearly one-half of all family farms in Brazil, captured only 5 percent of the overall *custeio* envelope and only 20 percent of *investimento* (Figures 3 and 4).



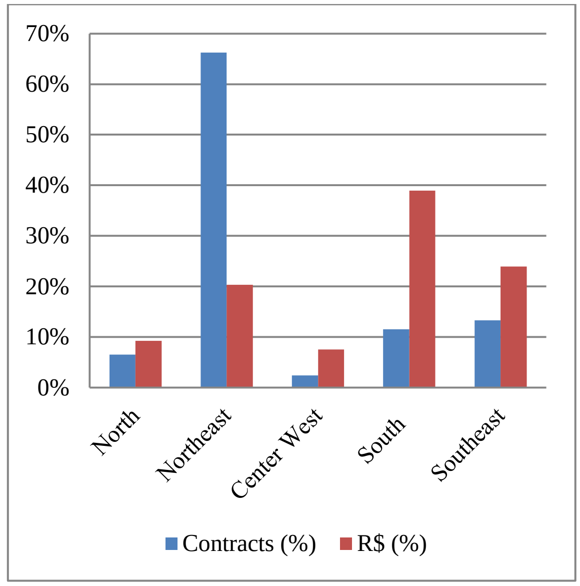


Figure 3: PRONAF Production Credit (2011)

Figure 4: PRONAF Investment Credit (2011)

Source: Banco Central do Brasil (2011)

Specific to *custeio*, the family farms of the South and Northeast macroregions stand out: in the South, for their ability to access PRONAF for their short-term production expenditures, whereas in the Northeast for their virtual absence from this line of finance. Furthermore, while *custeio* contracts in the South average R$9,821, in the Northeast the mean contract is less than one-half of this amount (R$4,406). Turning to *investimento*, mean contract values for the South and Northeast are R$26,000 and R$2,340, respectively. In the case of the Northeast, such levels of investment (at less than one-tenth of those for the South) are clearly inadequate to facilitate technology adoption.

Table 4: Share of PRONAF Resources by Income Classes

PRONAF Category	**2004**	**2007**
A	8	4
B	7	6
C	25	15
D	37	40
E	12	20
Others	11	15
TOTAL	**100**	**100**

Source: Guanziroli et al (2010)

Guanziroli et al (2010) assess the distribution of PRONAF resources across the various classes (Table 4). In this case, PRONAF-A is a special line of credit for land reform beneficiaries, while PRONAF-B through PRONAF-E correspond to groups defined by increasing levels of income. In 2004, the highest income levels (i.e., PRONAF-D and PRONAF-E) captured 49 percent of total resources while the two poorest levels (PRONAF-A and PRONAF-B) garnered 15 percent. This gap grew in 2007, with the two highest income levels having captured 60 percent of total resources, while the lowest income levels captured 10 percent.

1.6 Assessing Farm Productivity in Brazil

The productivity of family farms is an important determinant of their income, and of poverty per se in rural areas. In this research, we explore differences in productivity: (1) between two types of farms, namely family and non-family farms; and (2) among four different sizes of farms, as measured in hectares. Among the many determinants—geographical, technological, institutional, and idiosyncratic—we will pay particular attention to the role of public policies such as those that influence the use of credit (i.e., PRONAF), technical assistance (i.e., ATER), and irrigation.

To analyze the determinants of productivity among all farms in Brazil, the base model we estimate assumes that the relevant dimensions of technological heterogeneity are: 1) farm type (i.e., family versus non-family); and 2) farm size (i.e., four farm size classes in terms of hectares). The model allows the technology to vary across these eight groups. A second group of models is estimated just for family farms. Technological heterogeneity is considered by farm size and farm type which depends, for example, on the use of credit or technical assistance. For all these models, we assume that farm size is exogenous and related to the stock of wealth of a farmer, which only changes slowly over time. It is not unreasonable to assume that wealth is fixed for a single cross-section of data. Land markets, in addition, are rather thin in Brazil, which provides additional justification for treating farm size as fixed.

Once we condition on farm size, the distinction between farm types, such as family and non-family farms, is endogenous, meaning that hiring of labor, for example, is a choice variable, and if a farm uses more hired labor than family labor it is classified as a non-family farm.[4]

[4] Rather than ignore this "form endogeneity", and run the risk of producing coefficient estimates that are biased due to selectivity, we use the Heckman two-stage approach to control for unobservables that influence this decision. First, we estimate a probit for the probability of choosing to be a family farmer, and then separately estimate production functions (that incorporate farm size heterogeneity) for family and non-family farms. The inverse Mills ratio, which is calculated from the probit, is included in the second stage to control for unobservables that might

By specifying the model as above, i.e., an endogenous choice that is conditional on farm size, we can examine a variety of additional questions within the same framework. For example, conditional on farm size, we can examine how the endogenous choice to use credit, technical assistance, or irrigation leads to heterogeneous technologies across these dimensions.

Box 3: A Word about "Representative Farms"

The data we use are aggregated into representative—or average—farms at the municipal level. There are thousands of them in each of Brazil's five macroregions. In the base model, for example, the data are aggregated so that there are eight representative farms in each municipality. Brazil had over 5,500 municipalities in 2006. The eight average farms correspond to the four farm size classes for both family and non-family farms in each location. The aggregation was necessary so that we could work with the data outside of IBGE's premises without jeopardizing its confidentiality. Aggregation implies that we explore variation *between the representative farms*, but not *variation within them*.

The model also permits us to conduct counterfactual simulations that address what would happen to the productivity and income of family farms under a variety of scenarios. Here, we conduct two experiments. The first decomposes the differences in productivity across the eight types of farms into components that are attributable to differences in technology, the intensity of input usage, and the non-factor characteristics that influence outcomes, such as: (1) environmental factors; (2) distance to markets; (3) education; (4) experience; and (5) access to public goods and policies. The second is conducted for all "representative farms" in the data (Box 3). For each of these farms, we simulate the productivity gain of using a different technology. The simulations can be conducted unconditionally, or conditional on farm size or type. The results include average productivity gains for each group, as well as the distribution of the choices that optimize the gains.

Section 2 of this Report discusses descriptive statistics on family and non-family farms in order to provide context for the model and estimation results. In Section 3 we provide details on the model specification, namely the heterogeneous translog production function that will be used to estimate productivity, and discuss the data and the construction of the variables. Section 4 presents the initial results, Section 5 extends the model and Section 6 presents conclusions and policy implications.

---

have influenced both selection and production. Heckman's approach is appropriate for linear models. See Section 3 for more details.

2. Understanding the Family Farm in Brazil: Findings from the 2006 Agricultural Census

Family farms represent the overwhelming majority of farm establishments in Brazil. Yet family farms tend to be small, and lack physical and human capital. The combination of these factors contributes to low income levels and high poverty rates. We present information in this section that highlights these issues. The section is divided into four parts. The first part describes the relative importance of family farms in Brazil in terms of number of farms, area, and output. It also shows the distribution of these variables by farm size. The second part explores poverty among farms in Brazil, and shows how it differs for family and non-family farms. Evidence is also presented on the relationship between poverty, farm size, and land productivity. Part three seeks to question the view that family farms are more productive than non-family farms. While true in the aggregate, we argue that this is an artifact of differences in the size distribution of the two groups. Small farms have higher land productivity than large farms, and family farms tend to be small. When one compares farms of the same size, non-family farms uniformly have higher land productivity. This appears to be a result of a more intensive use of inputs and physical capital, not of an inherent superiority of one type over another. The final part of this section presents information on key characteristics of producers and their farms. It highlights the low levels of human and physical capital, as well as scant use of modern inputs, among a large share of farms – particularly family farms – in Brazil.

The data in this section are from the 2006 Agricultural Census. We present averages and cross tabulations at high levels of aggregation. Because we wanted to be able to match official data available on IBGE's website (www.ibge.gov.br) and in published volumes of the Census, no attempt was made to first clean the data. Most of the data used subsequently to estimate the econometric models found in this Report are also drawn from the 2006 Agricultural Census. Those data were filtered and cleaned prior to constructing the variables that are used for model estimation.[5]

[5] Thus, descriptive statistics for these filtered data may differ from the published IBGE data.

2.1 Size, Distribution and Geographic Dispersion of Family Farms

Of the 5.175 million farm establishments in Brazil in 2006, 84 percent are classified as family farms according the legal definition.[6] Table 5 shows that the share of family farms varies considerably across Brazil's five macroregions. Family farms represent as little as 69 percent of the farms in the Center-West, and as much as 89 percent in the Northeast. Although they are 84 percent of all farms, family farms only possess 24 percent of the land in agricultural establishments in Brazil. The share of land in family farms is as low as 9 percent in the Center-West and as high as 37 percent in the Northeast. The Northeast is home to 50 percent of all family farms nationwide, with another 19 percent residing in the South. The econometric analysis in Section 4 will focus on these latter two regions.[7]

Figure 5 shows the size distribution of family farms in the Northeast and South. The extremely small size of a significant portion of farms in the Northeast highlights part of the reason why there is so much rural poverty in this macroregion. Aside from the "producers without area,"[8] who have zero land and are excluded from the econometric analysis below, 20 percent of family farms have less than one hectare (ha) of land, and 33 percent have between one and five ha.

[6] The unit of analysis in the Agricultural Census is an establishment. We use farm and establishment interchangeably in this report. For the definition of an establishment, see IBGE (2009) and Helfand and Brunstein (2001).

[7] In order to reduce the heterogeneity of the analysis of the Northeast, and focus more squarely on poverty, the econometric analysis is limited to the semi-arid portion of the Northeast, thereby excluding the humid coastal areas (*zona da mata*), the transition between the two (*agreste*), and the transition to the Amazon rainforest (*meio norte*). The semi-arid region also excludes a piece of the *cerrado* grain belt which borders the Center-West. The official definition of the semi-arid region includes a small portion of Minas Gerais, in the Southeast region of the country, which borders the Northeastern state of Bahia. The municipalities in Minas Gerais were not included in the econometric analysis. The semi-arid region has 70 percent of the farms in the Northeast, and has the same share of family farms (89 percent) as the broader region.

[8] This is a new category of establishment that did not exist in previous Agricultural Censuses. IBGE (2009, p. 32), defines "producers without area" as employees of a farm who produce agricultural output of their own, under their own administration. If their production is supervised by the landowner, it counts as part of the output of the landowner's establishment. There were 255,000 establishments of this type in 2006. This is likely to be one of the reasons why there was an increase in the total number of establishments between 1995-96 and 2006.

Table 5: Number and Area of Establishments by Family Farm Type and Region

Region	Type	Establishments	Share of region	Area (ha)	Share of region	Brazil Share Estab. F	Brazil Share Estab. NF	Brazil Share Area F	Brazil Share Area NF
Brazil	Total	5,175,489		329,941,393					
	NF	807,587	0.16	249,690,940	0.76		1.00		1.00
	F	4,367,902	0.84	80,250,453	0.24	1.00		1.00	
North	Total	475,775		54,787,297					
	NF	62,674	0.13	38,139,968	0.70		0.08		0.15
	F	413,101	0.87	16,647,328	0.30	0.09		0.21	
Northeast	Total	2,454,006		75,594,442					
	NF	266,711	0.11	47,261,842	0.63		0.33		0.19
	F	2,187,295	0.89	28,332,599	0.37	0.50		0.35	
Southeast	Total	922,049		54,236,169					
	NF	222,071	0.24	41,447,150	0.76		0.27		0.17
	F	699,978	0.76	12,789,019	0.24	0.16		0.16	
South	Total	1,006,181		41,526,157					
	NF	156,184	0.16	28,459,566	0.69		0.19		0.11
	F	849,997	0.84	13,066,591	0.31	0.19		0.16	
Center-west	Total	317,478		103,797,329					
	NF	99,947	0.31	94,382,413	0.91		0.12		0.38
	F	217,531	0.69	9,414,915	0.09	0.05		0.12	

Source: Agricultural Census-IBGE (2006)
Note: F=family, NF=non-family.

Another 33 percent have between five and 50 ha. In the South, by contrast, 69 percent of family farms have between five and 50 ha. Thus, although the overwhelming majority of family farms in both regions have less than 50 ha, in the Northeast they are concentrated between 0-5 ha, while in the South there are many more between 10-50 ha. Given the earlier assumption of wealth exogeneity, this finding has important implications for poverty across the macroregions.

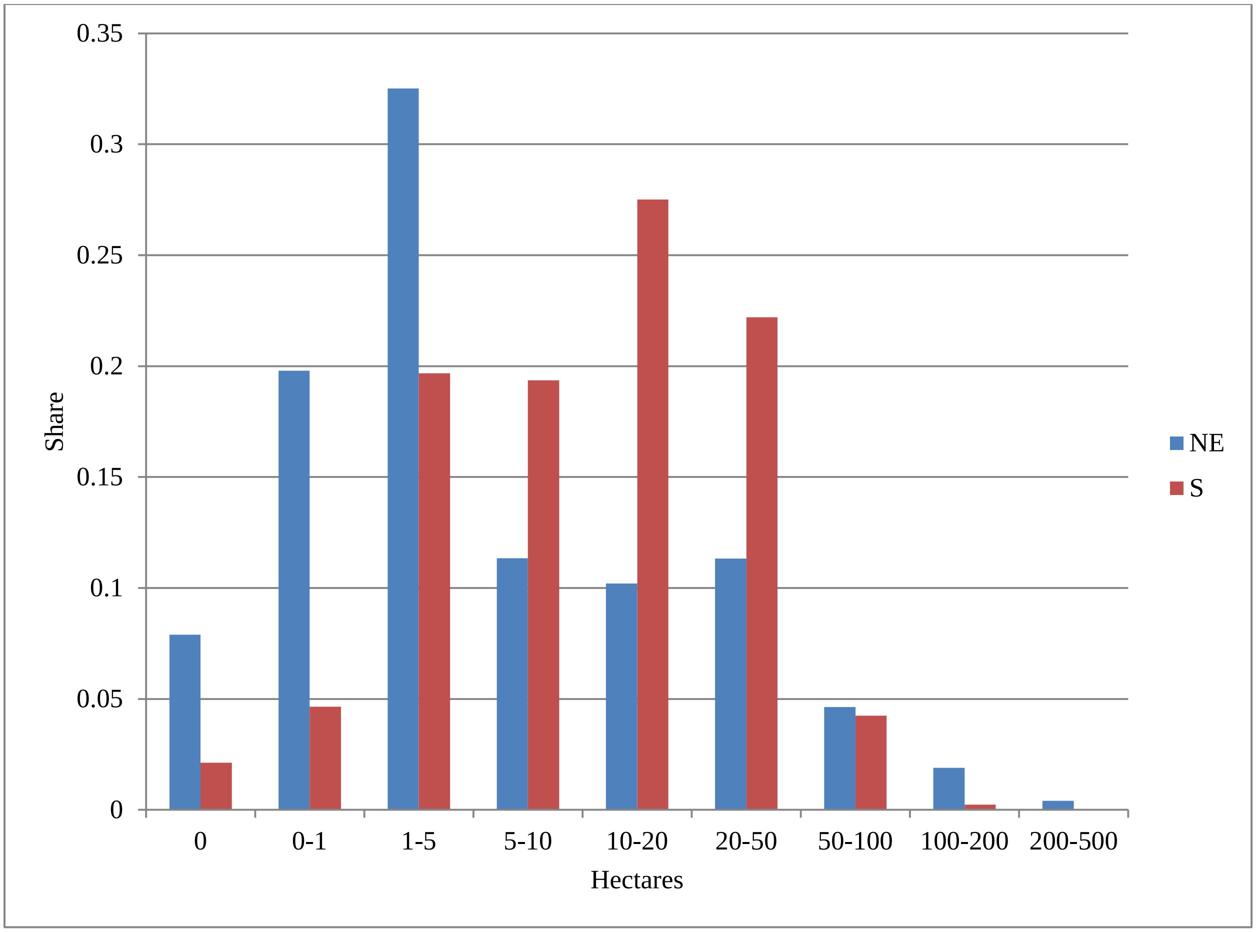


Figure 5: Distribution of Family Farms by Size: Northeast and South

Source: Agricultural Census-IBGE (2006)

Figure 6 shows the distribution of value of output for family farms according to farm size in hectares. The data indicate a greater proportion of relatively larger size family farms in the South relative to the Northeast. In part as a result of the differences in the size distribution of farms, farms between 10ha and 100ha in the South account for 65 percent of output, whereas farms of this size only produce 41 percent of output in the Northeast. Farms between 1-5ha comprise 27 percent of output in the Northeast.

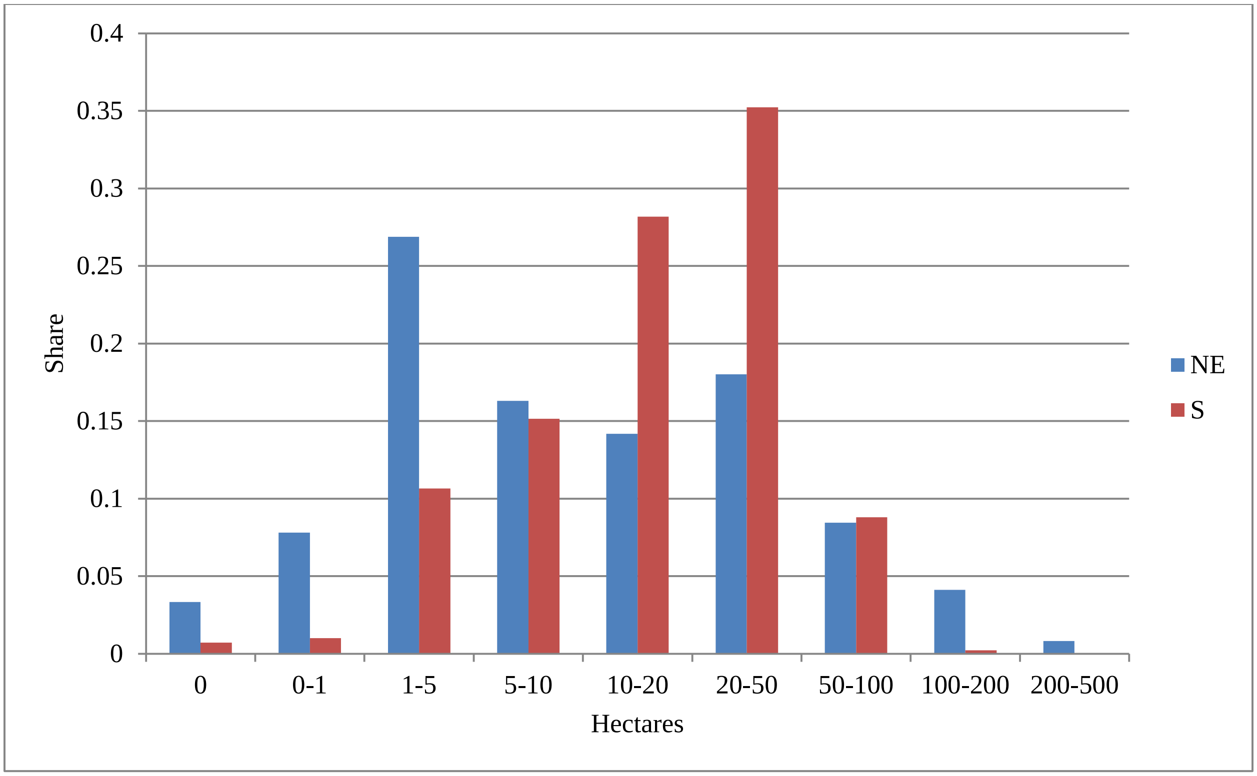


Figure 6: Distribution of Family Farm Value of Output by Size, Northeast and South

Source: Agricultural Census-IBGE (2006)

2.2 Family Farms and Poverty

Helfand et al. (2011) used the 1995-96 Agricultural Census to study poverty among agricultural producers in Brazil. It is not common to use the Agricultural Census for this purpose because the unit of analysis is the *establishment* rather than the *household*. The authors showed, however, that at the national level, the poverty rate was a little lower, but consistent with rural poverty rates based on the Demographic Census. The ranking of poverty across regions was the same, and the municipal level correlation between agricultural poverty (according to the 2006 Agricultural Census) and rural poverty (according to the Demographic Census) was 0.80. For this reason, we believe that the Agricultural Census can be used to study poverty among agricultural producers. Relative to 1995-96, the questionnaire used for the 2006 Census had a number of improvements that make poverty analysis even more feasible. For example, the 2006 Agricultural Census gathered some information on the household and on income earned outside of agriculture. The methodology used to calculate poverty is described in the next paragraph.

Table 6: Poverty Measures – Northeast Brazil

Income sources	Family size	Poverty (%)		Extreme Poverty (%)	
		SR	LR	SR	LR
Production	AE	72	75	62	65
Prod. + labor	AE	69			
Prod. + labor + transfers	AE	56	58	43	47
Production	N	76			
Prod. + labor + transfers	N	61	63	48	51

Source: Agricultural Census-IBGE (2006)

Notes:

Family size: AE=adult equivalent full time workers; N= all family members.

SR= short run (only variable costs are deducted from value of output).

LR=long run (estimated depreciation of assets is added to variable costs to obtain total costs).

Poverty lines: 1/2 and 1/4 minimum wage of August 2000 converted to 12/2006 values.

Table 6 presents a variety of poverty measures for the Northeast macroregion in order to illustrate what can be measured with the 2006 Agricultural Census. The first column defines the sources of income included in each measure. Three possibilities are contrasted: (1) income solely from agricultural production; (2) income from production and other earned income e.g. off-farm labor; and (3) income from these two sources plus government transfers (including social security and conditional cash transfers such as *Bolsa Familia*). The second column shows that family size can be measured as adult-equivalent full-time family labor involved in farm production (AE), or by the total number of family members who were involved in production (N). The next two columns calculate poverty rates with a poverty line that is commonly used in Brazil (½ minimum wage per capita from August 2000), while the final two columns use an extreme poverty line of ¼ minimum wage per capita. These are the same poverty lines, adjusted for inflation, that were used in Helfand et al. (2011). Agricultural profit is calculated with the value of output, rather than sales, in order to account for non-monetary income. Both poverty and extreme poverty can be calculated with different measures of costs deducted from the value of agricultural output. Columns 3 and 5 use a measure of short-run variable costs (monetary expenditures), while columns 4 and 6 also impute a long-run value for depreciation of buildings, machines, trees, and cattle.

The third column of Table 6 shows a short run (SR) measure of poverty that only deducts variable costs from the value of output. In the Northeast, 72 percent of producers did not generate sufficient short-run farm profits to raise the adult-equivalent full-time equivalent family labor above the poverty line. When off-farm earnings are added to income, poverty only falls to 69 percent for these farms. Transfers (e.g., *Bolsa Família*) make much more of a difference, lowering the rate of poverty to 56 percent. When family members are simply counted (rows 4 and 5 of Table 6), with no adjustments for days worked or the demographic composition of the family labor, poverty is four or five percentage points higher, depending on the measure of income that is used. When depreciation costs are included in order to estimate a long run (LR) measure of profits, poverty only increases by two or three percentage points (column 4). Columns 5 and 6 show that most poverty among agricultural producers is rather deep. The first row of Table 6 shows that when LR costs are included 75 percent of producers in Northeast are poor. When a much lower poverty line was applied (i.e., 1/4 rather than 1/2 a minimum wage per capita), 65 percent were still considered poor. Subsequent tables and figures use the ½ minimum wage poverty line, short run costs, and adult equivalent family labor.

Table 7 shows how poverty rates in the Northeast vary by farm type and size. For family farms with greater than zero and less than five ha of land, 81 percent were poor. Poverty drops by 14 percentage points, (to 67 percent) for farms with 5 to 20ha of land, and by another 7 percentage points (to 60 percent) for farms with 20-100ha. Off-farm income was not very important for these farms, as the poverty rate only drops by two percentage points in all cases when this income is included. The inclusion of transfers--both CCTs and pensions--is much more significant. For farms with 0-

Table 7: Poverty by Farm Size, Farm Type, and Income Source (Northeast)

Type/Income sources	Poverty (%) by Farm Size (ha)			
	0-5	5-20	20-100	100-500
Family				
Production	81	67	60	54
Prod. + labor	79	65	58	52
Prod. + labor + transfers	65	51	44	41
Non-family				
Production	81	62	51	51
Prod. + labor	53	46	42	47
Prod. + labor + transfers	44	39	37	43

Source: Agricultural Census-IBGE (2006)

Note: Relative to Table 6, this table only uses the short-run poverty measure with adult equivalent labor and the ½ minimum wage poverty line.

5ha, poverty falls by 14 percentage points (to 65 percent) when transfers are included. Thus, the impact of transfers is roughly of the same magnitude as the impact of increasing farm size from 0-5ha to 5-20ha. In the first case, however, poverty is reduced through a continuous flow of transfers, while in the second case capital accumulation in the form of land permits generating a higher stream of earned income.

Non-family farms with 0-5ha had the same rate of poverty as family farms (81 percent) when only agricultural income was considered. By contrast, off-farm income was very important for this group: its inclusion lowered poverty by 28 percentage points. Thus, small non-family farmers might not only hire labor in (one of the criteria for distinguishing between family and non-family farms), but also hire a significant amount of their own labor out. Transfers were also important for this group, lowering poverty by another 9 percentage points. Thus, the most important difference--in so far as poverty is concerned--between family and non-family farms that have only 0-5ha was off-farm income. *This is a novel and significant finding*. It also applies more generally to the two groups. In all of Brazil, 45 percent of non-family farmers had off-farm work, whereas only 26 percent of family farmers did. It is possible that non-family farmers have better labor market opportunities--perhaps due to education, skills, or geography--that generate cash which in turn permits them to hire labor to work on their own farms. This is a hypothesis that should be explored in future research.

Off-farm income becomes less important for non-family farms in the Northeast as size increases. When total income is considered, non-family farms with 100-500ha appear to have slightly more poverty than family farms of the same size and more poverty than smaller non-family farms.

Table 8 presents data on poverty by size and type in the South macroregion. Based on agricultural income alone, family farms with 0-5ha of land had 19 percentage points less poverty than in the Northeast. Off-farm income is a little more important for this group in the South, and transfers play a similar role in both regions. Thus, the reason why family farms with 0-5ha in the South had 22 percentage points less poverty than in the Northeast is almost entirely due to differences in agricultural income. Farms in the South, as we shall see below, are much more productive.

Non-family farms with 0-5ha in the South macroregion had half the poverty that family farms did when all income sources are considered. Worthy of note is that only 22 percent of these farms were poor. Although they generated a little more farm income than comparably sized family farms, the main difference was that non-family farms earned much more off-farm income. Similar to what was observed in the Northeast macroregion, the inclusion of off-farm income lowers the poverty rate by 32 percentage points. The importance of off-farm income declines as farm size grows, and is not fully compensated for by farm income.[9]

Table 8: Poverty by Farm Size, Farm Type and Income Source (South)

Type/Income sources	Poverty (%) by Farm Size (ha)			
	0-5	5-20	20-100	100-500
Family				
Production	62	35	30	-
Prod. + labor	57	33	28	-
Prod. + labor + transfers	43	26	23	-
Non-family				
Production	58	37	38	51
Prod. + labor	26	23	33	47
Prod. + labor + transfers	22	21	30	44

Source: Agricultural Census-IBGE (2006)

Note: Relative to Table 6, this table only uses the short run poverty measure with adult equivalent labor and the ½ minimum wage poverty line.

Figures 7 and 8 show the relationship between poverty, farm size and land productivity for all farms in the Northeast and South macroregions. The figures are based on farm income alone, and divide producers of each size into quintiles of land productivity. Thus, category 1 corresponds to the 20 percent of producers with the lowest land productivity, and category 5 corresponds to the 20 percent with the highest land productivity for each farm size class. Figure 7 for the Northeast macroregion illustrates three important points.

- First, land distribution is a significant obstacle to poverty reduction. As we saw in Table 5 and Figure 5, some 89 percent of farm establishments in the Northeast are family farms, more than half of which have between 0-5ha. Even if all of these farms could raise land productivity to the level of the 5th quintile, most would not escape poverty. Over 70 percent of the existing farms of this size, and level of productivity, remain poor.

[9] The fact that non-family farms with 20-100ha appear to have more poverty than smaller non-family farms could reflect the fact that some of these larger farms are unproductive, or that the Census did not fully capture the off-farm earnings of these farmers.

- Second, if poverty reduction is to take place based on agriculture income, it will require both larger farm sizes and higher levels of productivity in the Northeast. The two must go together. In order to bring poverty down to around 50 percent, one needs either 5-20ha of land operating in the top quintile of productivity, or 20-100ha in the 3rd quintile.

- Finally, farm income alone is unlikely to solve the problem of poverty for agricultural producers in this region. More than 40 percent of farms with 20-100ha in the top quintile of productivity were poor when only farm income was considered, and the overwhelming majority of producers have neither this amount of land nor this level of productivity. Multiple policy approaches are necessary, including land reform, technical change, enhancing off-farm income opportunities, education, and transfers.

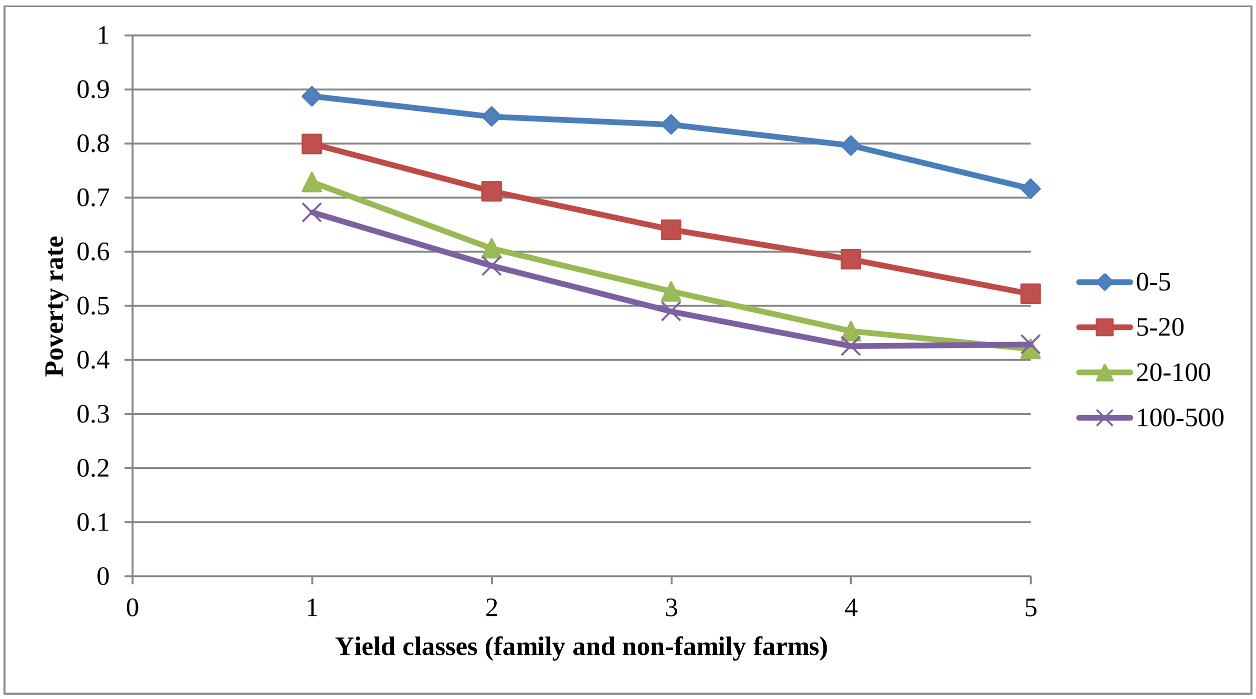


Figure 7: Poverty by Farm Size and Yield (Northeast)

Source: Agricultural Census-IBGE (2006)

Productivity plays a more important role in the South, both in terms of the average level and the potential gains. For farms with 0-5ha in the South, Figure 8 shows that poverty falls from 77 percent to 47 percent as land productivity increases from the bottom to the top quintile. And for farms with 5-20ha, poverty falls to 33 percent for farms in the 3rd quintile of the productivity distribution. While agriculture is also not the only pathway out of poverty in the South, it can play a much more important role than in the Northeast. It appears to be more like a highway, and less like a narrow pathway, in this region of Brazil.

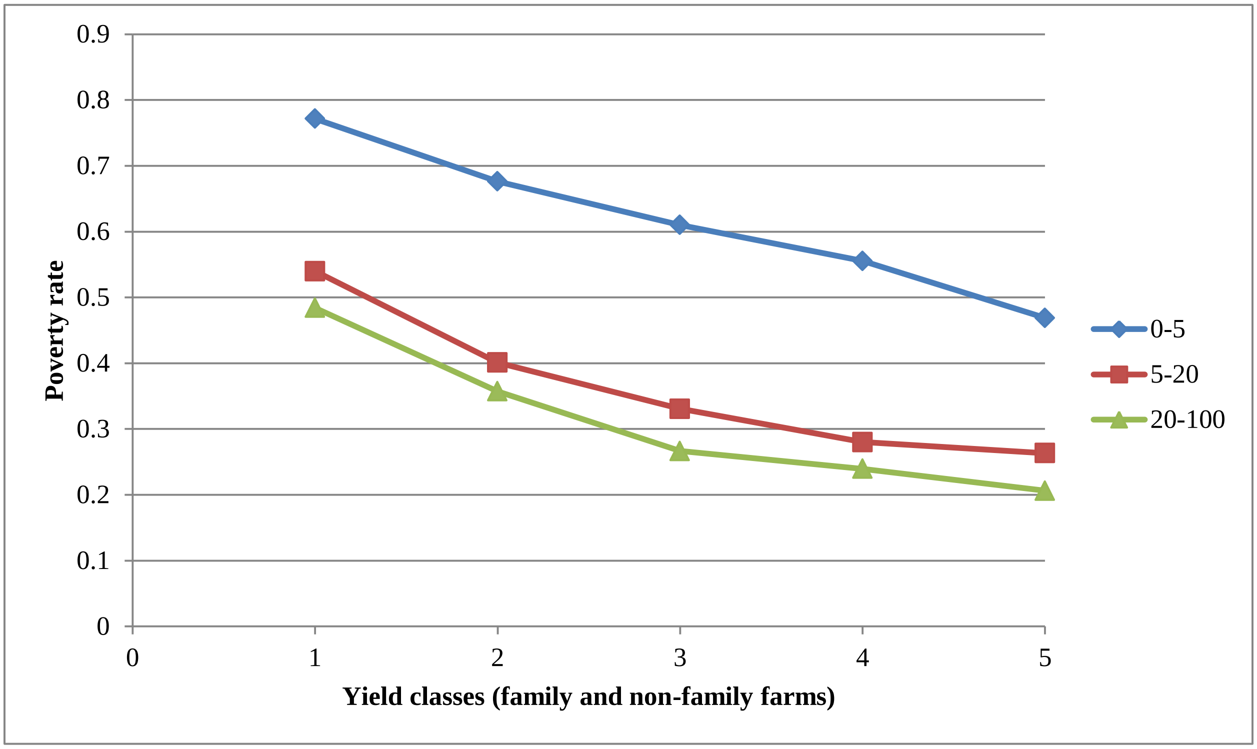


Figure 8: Poverty by Farm Size and Yield (South)

Source: Agricultural Census-IBGE (2006)

2.3 Family Farms and Land Productivity

The discussion above highlighted the potential importance of productivity improvements for poverty reduction. We now address the relationship between land productivity and type of farm (i.e., family and non-family). In particular, we question the view that family farms are more productive than non-family farms. While true on average, we believe that this is a reflection of differences in the composition of farm sizes. It is a well-established fact that small farms have higher land productivity than large farms. Because family farms comprise a much higher proportion of small farms, this is the reason why they have higher average land productivity. But when one controls for farm size, this superiority disappears.

Table 9 shows value of output per hectare by farm type and size for all of Brazil. On average, land productivity of family farms (R$648) is about 80 percent higher than land productivity of non-family farms (R$354). Although Guanziroli et al. (2010) use a slightly different definition of family farm, which includes medium size farms in the family farm category, their data are quite similar. They show a comparison of R$530 for family farms versus R$385 for non-family farms in 2006, and a nearly identical family farm versus non-family ratio of productivities in 1995-96. When family and non-family farms are compared conditional on farm size, however, the results are reversed (see Column 3 of Table 9). In every farm size category, non-family farms produce more value of output per hectare than family farms. The same result holds within specific regions of the country such as in the South or the Semi-Arid portion of the Northeast. In our view, this challenges the notion that family farms are more productive because of the form in which these farms are administered. While a family administration might be better, and incentives for family members might be superior, there appear to be other factors that matter more. More specifically, at every farm size, non-family farms use more purchased inputs and capital per hectare, which more than compensates for the use of less family labor.

Table 9: Value of Output per Hectare by Farm Type and Farm Size (Brazil)

Type / Size(ha)	Mean Yield	Ratio: F/NF
Family (F)	**648**	1.83
0-1	7,192	0.54
1-5	2,858	0.39
5-10	1,867	0.44
10-20	1,198	0.43
20-50	668	0.36
50-100	308	0.30
100-200	119	0.17
200-500	82	0.23
Non-Family (NF)	**354**	
0-1	13,263	
1-5	7,315	
5-10	4,224	
10-20	2,794	
20-50	1,877	
50-100	1,019	
100-200	685	
200-500	352	
500-1000	260	

Source: Agricultural Census-IBGE (2006)

Note: the mean for each type excludes farms with zero land.

By way of example, we briefly discuss several variables that were constructed for the econometric analysis (presented later in this Report). In the Semi-Arid Northeast, non-family farms with 5 to 20 hectares of land, when compared with family farms of similar holdings: (i) spend 45 percent more on purchased inputs per hectare; and (ii) have 60 percent more capital per hectare. In the South, non-family farms of this size spend 27 percent more on purchased inputs and have 66 percent more capital per hectare than similarly-sized family farms. In both regions, the difference in the use of full-time equivalent family labor per hectare is less than 10 percent for farms of this size. These differences in input intensities help to explain the higher yields for non-family farms at each farm size, and suggest that many family farms operate in a low-level equilibrium.

2.4 Gender, Human Capital and Input Use on Family Farms

Table 10 presents key characteristics of family farmers by gender. Most family farmers are men between 35 and 65 years old with little formal education and more than ten years of experience running the current farm. More precisely, 86 percent of family farms are run by men, 41 percent of whom are either illiterate or have no formal schooling, and another 44 percent did not finish primary school. Only 7 percent completed high school or more. Two-thirds of these men are between 35 and 65 years old, with equal shares above and below these ages, and 62 percent have at least ten years of experience on the current farm. It is not clear if "experience" should be interpreted as a positive characteristic that contributes to human capital and productivity, or as a reflection of

Table 10: Gender of Family Farmer vs. Age, Education, and Experience

		Gender		
		Female	Male	Difference[1]
Share of total		0.14	0.86	
Education	Illiterate	0.33	0.26	0.07
	Literate, but no schooling	0.15	0.15	0.00
	Primary incomplete	0.38	0.44	-0.06
	Primary complete	0.07	0.08	-0.01
	Secondary complete	0.05	0.05	0.01
	Higher education	0.03	0.02	0.00
	Total	1.00	1.00	
Age	0-25	0.03	0.03	0.00
	25-35	0.11	0.14	-0.03
	35-45	0.18	0.22	-0.04
	45-55	0.22	0.23	-0.01
	55-65	0.23	0.20	0.02
	65-	0.24	0.17	0.06
	Total	1.00	1.00	
Years in charge of farm	0-1	0.02	0.03	0.00
	1-5	0.16	0.17	-0.02
	5-10	0.16	0.18	-0.02
	10-	0.66	0.62	0.04
	Total	1.00	1.00	

Source: Agricultural Census-IBGE (2006)

[1] Likelihood of female, compared to male, with a given characteristic (in % difference).

being "trapped" in a low-productivity activity. Low levels of education are a historical legacy in rural Brazil, particularly in the semi-arid Northeast, and are a persistent obstacle to reducing poverty. But more recent cohorts have higher levels of schooling: according to the 2006 household survey (PNAD), 20-24 year olds in rural areas now complete close to eight years of schooling on average (Helfand and Pereira, 2011).

Female family farmers are not much different than their male counterparts. They are 7 percentage points more likely to be illiterate, but roughly the same share has primary education incomplete or less, and they are some 8 percentage points more likely to be at least 55 years old.

Relative to family farmers, non-family farmers are quite different. They are 6 percentage points less likely to be over 55 years of age, 7 percentage points less likely to have at least 10 years of experience on the current farm, and 23 percentage points more likely to have at least a complete primary education. Figure 9 contrasts the education profiles of family and non-family farmers. It shows that 12 percent of non-family farmers have completed primary school, 12 percent have completed secondary school, and 14 percent have some form of higher education. Thus, as a group, they are somewhat younger and considerably more educated than the family farmers. Non-family farmers are also 15 percentage points more likely to have obtained their land through purchase, with 71 percent stating that this was the form in which they acquired their land.

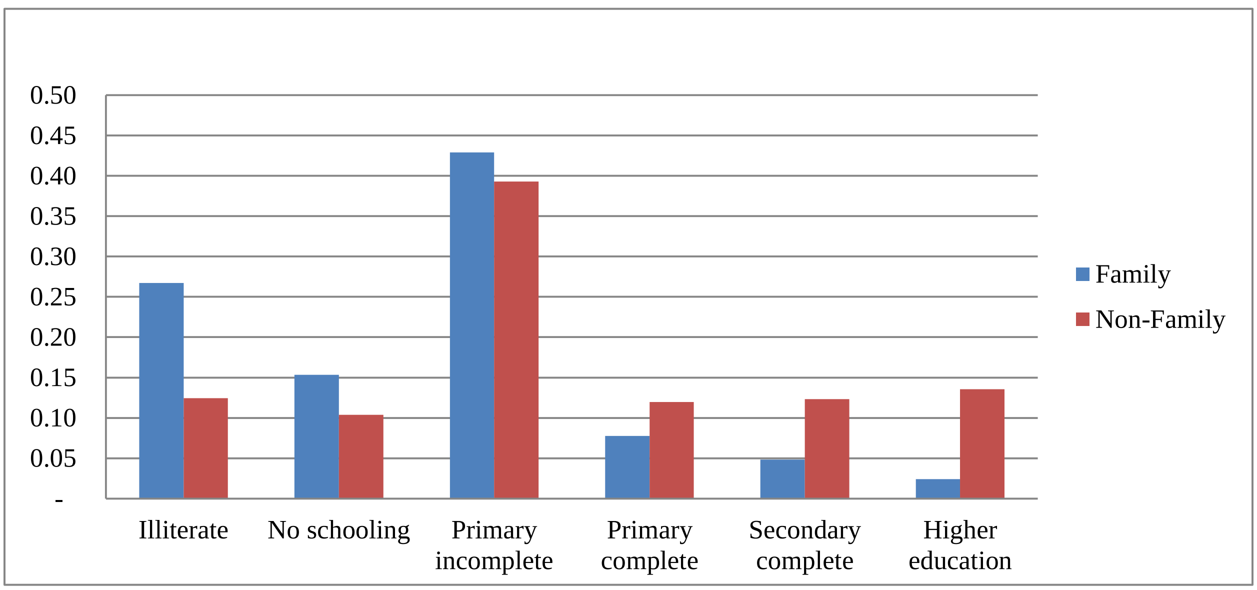


Figure 9: Schooling Among Agricultural Producers: Family vs. Non-Family

Source: Agricultural Census-IBGE (2006)

Table 11 shows the relationship between education and certain characteristics of production for family farms. Family farmers who have a secondary education or more are close to 10 percentage points more likely to be very highly specialized in a single crop. This is likely a reflection of increased knowledge and ability to bear risk. Around 80 percent of family farmers who are illiterate or have no schooling report not using either credit or technical assistance. When further pressed as to why they are not using credit, one-half of these family farmers state, "It was not needed." This share falls to around two-thirds for family farmers with a primary or secondary education, and to one-third for family farmers with additional schooling. Seven percent of family farmers use credit with technical assistance, while 11 percent use credit without technical assistance. The more educated family farmers are more likely to use credit and technical assistance together, and much more likely to use technical assistance, even without credit. Sixty-six percent of family farmers with higher education use technical assistance, either with or without credit, while less than 30 percent of farmers in the other educational categories do so. Of those family farmers who are either illiterate or have no formal schooling, close to 90 percent reported no use of technical assistance; this has important implications for the adoption of new technologies and farming practices required to boost productivity.

The use of fertilizers is also quite different between those farmers who are illiterate or have no schooling and the other groups. Over 75 percent of the less educated farmers report not using fertilizers, while this share falls to around 56 percent for all other levels of education. The use of irrigation rises with education, but never surpasses 9 percent of family farmers at any level of education.

The final two variables in Table 11 show productivity and income differences across levels of education. Mirroring the use of technical assistance and fertilizers, the bottom row shows that land productivity is significantly lower for family farmers who are illiterate or have no schooling. These are also likely to be the oldest farmers, with around half of each of these groups having farmers who are over 55. Among illiterate farmers, 84 percent have total farm income between zero and R$16,000 (i.e., PRONAF income classes B and C). This share falls to 63 percent for farmers with an incomplete primary education, and continues to decline steadily until reaching 54 percent for family farmers with higher education. There appears to be a discrete difference between those farmers who are illiterate or have no schooling and the rest.

Above this level, education appears to lower the probability of having an income below R$16,000 and increase the probability of having an income above R$80,000. There is no doubt that education matters for generating income.

Table 11: Education vs. Other Characteristics of Family Farmers

Characteristics	Illiterate	Literate, no schooling	Primary Incomplete	Primary Complete	Secondary Complete	Higher Education	Total
All Family Farmers	0.27	0.15	0.43	0.08	0.05	0.02	1.00
Degree of specialization[a]							
Very high	0.18	0.21	0.20	0.26	0.28	0.29	0.20
High	0.34	0.34	0.36	0.34	0.30	0.27	0.35
Moderate	0.35	0.30	0.30	0.24	0.19	0.16	0.30
Low	0.05	0.04	0.04	0.02	0.02	0.01	0.04
n.a.	0.08	0.12	0.10	0.14	0.21	0.26	0.11
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Use of Credit (C) and Technical Assistance (TA)							
C, TA	0.02	0.03	0.10	0.11	0.10	0.13	0.07
C, no TA	0.12	0.11	0.11	0.09	0.08	0.02	0.11
no C, TA	0.05	0.08	0.14	0.17	0.20	0.53	0.12
no C, no TA	0.82	0.78	0.64	0.63	0.62	0.32	0.70
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Use of Fertilizers							
Purchased	0.12	0.18	0.36	0.36	0.34	0.35	0.27
Own	0.04	0.05	0.04	0.05	0.05	0.06	0.05
Yes, but not in 2006	0.02	0.02	0.03	0.03	0.04	0.04	0.03
No	0.82	0.75	0.57	0.56	0.57	0.55	0.66
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Irrigation							
Yes	0.04	0.05	0.06	0.08	0.09	0.09	0.06
No	0.96	0.95	0.94	0.92	0.91	0.91	0.94
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00
PRONAF Income Classes							
A (land reform)	0.02	0.02	0.03	0.04	0.06	0.07	0.03
B (< R$3000)	0.09	0.10	0.10	0.09	0.07	0.05	0.09
C (R$3000- R$16,000)	0.75	0.67	0.53	0.51	0.51	0.49	0.61
D (R$16,000 - R$45,000)	0.10	0.15	0.22	0.22	0.22	0.23	0.18
E (R$45,000 - R$80,000)	0.02	0.04	0.08	0.09	0.08	0.08	0.06
Not PRONAF	0.01	0.02	0.04	0.05	0.06	0.08	0.03
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Land Productivity (index)	1.00	0.95	1.30	1.45	1.55	1.39	

Source: Agricultural Census-IBGE (2006)

[a] Degree of specialization based on share of principal product in value of output.
1)100%-very high, 2) 65% -100%-high, 3) 35% - 65%-moderate, 4) =<35%-low, 5) n.a. = not classified.

2.5 Physical Capital of Family Farms

Having discussed the obstacles to poverty reduction created by farm size in the Northeast, and low levels of human capital in general, we now turn to the level of physical or fixed capital used by Brazilian farmers. There is no doubt that the amount of fixed capital used in production is an important determinant of income among agricultural producers. Butzer et al. (2010), for example, show an extremely high correlation between agricultural GDP per worker and agricultural fixed capital per worker based on cross-country evidence for the period 1967-2003. In light of this evidence, what is the fixed capital base for Brazilian farms?

Table 12 shows the value of assets on Brazilian farms for selected states and farm sizes. For Brazil as a whole, the second row of the table shows that farms between 0-10ha—92 percent of which are family farms — accounted for 49 percent of the farms that reported the value of assets. The assets on these farms represented seven percent of the total value of assets on Brazilian farms. Fifty-six percent of the reported value for small farms was held in the form of land. Thus, farms with 0-10ha had an average of R$15,523 in non-land assets. If one generously assumes a 10 percent rate of return, and a family of four, this would generate an income stream equal to about R$32 per capita per month, or roughly one quarter of the poverty line that we used above.

Small farms in the South and Southeast, such as in the states of Santa Catarina and São Paulo, were in much better shape when compared to Brazil as a whole. In these states, small farms had an average of around R$50,000 in non-land assets. In the Northeast, where farms are smaller and human capital levels are lower, farms also utilize much less capital in production. In Alagoas and Pernambuco, farms only had an average of about R$5,000 and R$7,500 of capital. Again assuming a 10 percent rate of return and a family of four, this would generate a stream of income per person equivalent to less than 15 percent of the poverty line. Thus, the low levels of productivity among family farms in the Northeast result, in part, from low levels of physical and human capital. In terms of income and poverty, the low levels of productivity are compounded by the high share of farms in the region that have less than five hectares of land.

Table 12: Value of Assets on Brazilian Farms

		Establishments		Value of assets		Value of land	Land share	Assets / Estab. (without land)	Income p.c. per month (10% Rate of Return)
		#	%	R$1000	%	R$1000	%	R$	R$
Brazil	Total	5,090,960		1,238,572,593		874,206,542	71	71,571	149
	0-10ha	2,477,071	49	86,737,919	7	48,286,078	56	15,523	32
Santa Catarina	Total	192,795		59,452,816		34,325,077	58	130,334	272
	0-10ha	69,390	36	7,059,189	12	3,793,130	54	47,068	98
Sao Paulo	Total	227,024		191,774,691		148,568,403	77	190,316	396
	0-10ha	84,298	37	11,956,097	6	7,700,777	64	50,479	105
Matto Grosso	Total	64,764		90,983,839		67,454,336	74	363,311	757
do Sul	0-10ha	13,396	21	781,003	1	507,844	65	20,391	42
Pernambuco	Total	300,211		11,588,374		6,377,153	55	17,359	36
	0-10ha	208,110	69	3,234,560	28	1,668,507	52	7,525	16
Alagoas	Total	121,271		12,970,929		10,823,045	83	17,711	37
	0-10ha	95,791	79	1,894,151	15	1,425,889	75	4,888	10

Source: Agricultural Census-IBGE (2006)

3. Toward a Model of Production and Productivity of Family Farms

3.1 Brief Description of the Model

The productivity of family farms is an important determinant of their income, and of poverty in rural areas. Two sets of models are estimated in order to analyze these determinants of productivity.[10] A base model is estimated for all farms in Brazil. It assumes that the relevant dimensions of technological heterogeneity are: 1) farm type (i.e., family versus non-family); and 2) farm size (i.e., four farm size classes in terms of hectares). The model allows the technology to vary across these eight groups. It uses a translog production function and is estimated with a Heckman, two-stage approach that seeks to correct for the endogeneity of farm type. A second group of models is estimated just for family farms. In this case, technological heterogeneity is considered by farm size and family farm "type", where type is defined by the use of credit, technical assistance, irrigation, or specialization in production. As with the previous set of models, a translog production function is estimated with a two-stage approach. Once the model is estimated, we conduct counterfactual simulations to explore what would happen to family farms' productivity under a variety of scenarios.

We conduct two types of experiments. The first uses a Oaxaca-Blinder approach to decompose the differences in productivity between the mean of each type and size of producer into components attributable to differences in 1) technology, 2) the non-factor characteristics that influence outcomes, and 3) the intensity of input usage. The non-factor characteristics include environmental factors, distance to markets, education, experience, and access to public goods and policies. The sum of the technology and non-factor components can be interpreted as the difference in total factor productivity (TFP) between any two groups. The second experiment is conducted for all "representative farms". For each of these farms, we simulate the productivity gain of using a different technology. The simulations can be conducted unconditionally, or conditional on farm size or type. The results include average productivity gains for each group, as well as the distribution of the choices that optimize the gains. Box 4 provides definitions for the variables used in the model estimation.

[10] See Section 7 (Technical Annex) for greater detail on the econometric modeling and counterfactual simulations.

Box 4: The Heterogeneous Translog Production Function

We estimate a heterogeneous translog production function with the following functional form: $y = \alpha(t) + x\beta(t) + x'x\gamma(c) + z\delta(c) + v$, where the left-hand side, *y*, is defined as the log of the total value of output. The inputs (x_i) and non-factor variables (z) are defined below:

Inputs (x_i), where:

- x_1: Family labor measured as the log of adult-equivalent, full-time equivalent units.
- x_2: Capital stock measured as the log of the value of on-farm assets, (excluding land) and including buildings, machines, perennial trees or planted forests, breeding and work animals, etc.
- x_3: Purchased inputs measured as the log of the value of expenditures on all inputs including hired labor.
- x_4: Land measured as the log of total land excluding land in natural forests and land that is unusable.
  Note: Because constant returns to scale (CRS) is imposed, output (y) and the inputs (x_1, x_2, x_3) are normalized by the land variable (x_4). Thus, x_4 does not appear explicitly in the model.

Non-factor variables (z) that shift the production function:

Policy-related variables (measured as the share of farms within each representative farm) $w_{i,}$ where:

- w_1: Use of credit.
- w_2 Use of technical assistance.
- w_3: Use of electricity on the farm.
- w_4 Use of irrigation.

Climate, geography and transactions costs (g_i), where:

- g_1-g_{11}: Eleven municipal-level climate, soil, and slope variables.
- g_{12} -g_{14}: Three variables that capture transactions costs: (a) the municipal distance to São Paulo (the economic center of the country); (b) the municipal distance to the state capital; and (c) the average distance by farm size to the seat of the municipality in 1995-96.

Characteristics of the producer and family (p_i), where:

- p_1: Born in state: share of producers born in the state where production occurs.
- p_2-p_6: Education: share of producers within each of five levels of education.
- p_7-p_9: Years in charge of farm: share of producers within each of three levels of experience on the current farm.
- p_{10}-p_{14}: Age: share of producers in each of five age categories.
- p_{15}: Gender of producer.
- p_{16}-p_{17}: Share of adult family members who are illiterate, and who are unskilled, (and thus largely do manual labor).

3.2 The Data: Representative Farms as a Unit of Analysis

It is important to emphasize that the models discussed above were not estimated with farm-level data. Due to issues associated with the confidentiality of the microdata, all of the data used were aggregated by municipality, farm size in hectares (0-5, 5-20, 20-100 and 100-500), and class (family, non-family). Data aggregation implies that we assume homogeneity within each aggregate observation, for example family farms with five to ten hectares in the municipality of Viçosa, MG. We call these "representative farms." The econometric analysis will explore variance between representative farms, but due to the aggregation, cannot examine variance within them (see Box 3 for a detailed discussion of representative farms). This is an additional reason for emphasizing technological heterogeneity rather than idiosyncratic inefficiency, as the latter is specific to the individual farm.

Prior to aggregating the data into representative farms by municipality, size, and class, we filtered the data to remove spurious observations. When working with a preliminary dataset, we observed unbelievably large values of land productivity and other variables. For this reason, and after some sensitivity analysis, we elected to filter the data as follows. We defined five variables that were important for our study, and removed the largest one-half of one percent (0.5%) of the observations for each variable in each state. The variables, all defined per hectare, were: (i) value of output; (ii) value of expenditures on inputs; (iii) value of the capital stock; (iv) value of sales; and (v) total number of family members working on the farm. Thus, by way of example, in the state of Pernambuco, where there were 304,788 farms, we removed 5,094 observations (or 1.67 percent of the total) prior to aggregation. This was somewhat less than 2.5 percent (5 exclusions x .05 percent each) because the filters were applied simultaneously and some observations fell into more than one filter.

One of the limitations of working with aggregate data is that, in order to protect the confidentiality of the data, IBGE censors the value of a variable that has fewer than three observations reporting information. Thus, if there were five family farms between 100 and 500 hectares in the municipality of Água Branca, Alagoas, but only two of them report having family members who worked on the farm less than 60 days, the value of this variable would appear as missing. In order to reduce the loss of representative farms, we decided to impute data when we had additional information and in cases that we were reasonably comfortable that the imputation would not affect the results. Thus, because in the example above we know that each farm had to have had at least one person working on the farm less than 60 days, we imputed the lower bound of one person per farm. In other cases, for example, if "unusable land" was missing and we needed to subtract this component from total area in order to calculate net area, we imputed the average unusable share of total area for that farm type, size, and region. In no case did we impute values for key variables such as total value of output or total amount of land. Because cells with few farms are the ones that get censored, this implies that we lose a higher share of representative farms (cells) than actual farms.

The econometric analysis was conducted for the South macroregion and the Semi-Arid portion of the Northeast macroregion. Figure 10 displays a map of the Semi-Arid portion of the Northeast macroregion. Table 13 shows the number of representative farms, and individual farms, of each size and class in the Semi-Arid and South. We estimate the models in each region with around 7,000 observations (7,144 in the Semi-Arid and 6,821 in the South) which aggregate the information from 1.481 million farms in the Semi-Arid and 926,000 farms in the South, which together account for nearly 50 percent of all farm establishments nationwide. On average, there are around 45 farms in each representative non-family farm in the Semi-Arid, and around 38 in the South. Representative family farms have around 350 and 240 farms contained in each observation in the Semi-Arid and South, respectively, although there are far fewer in the 100-500 hectare class.

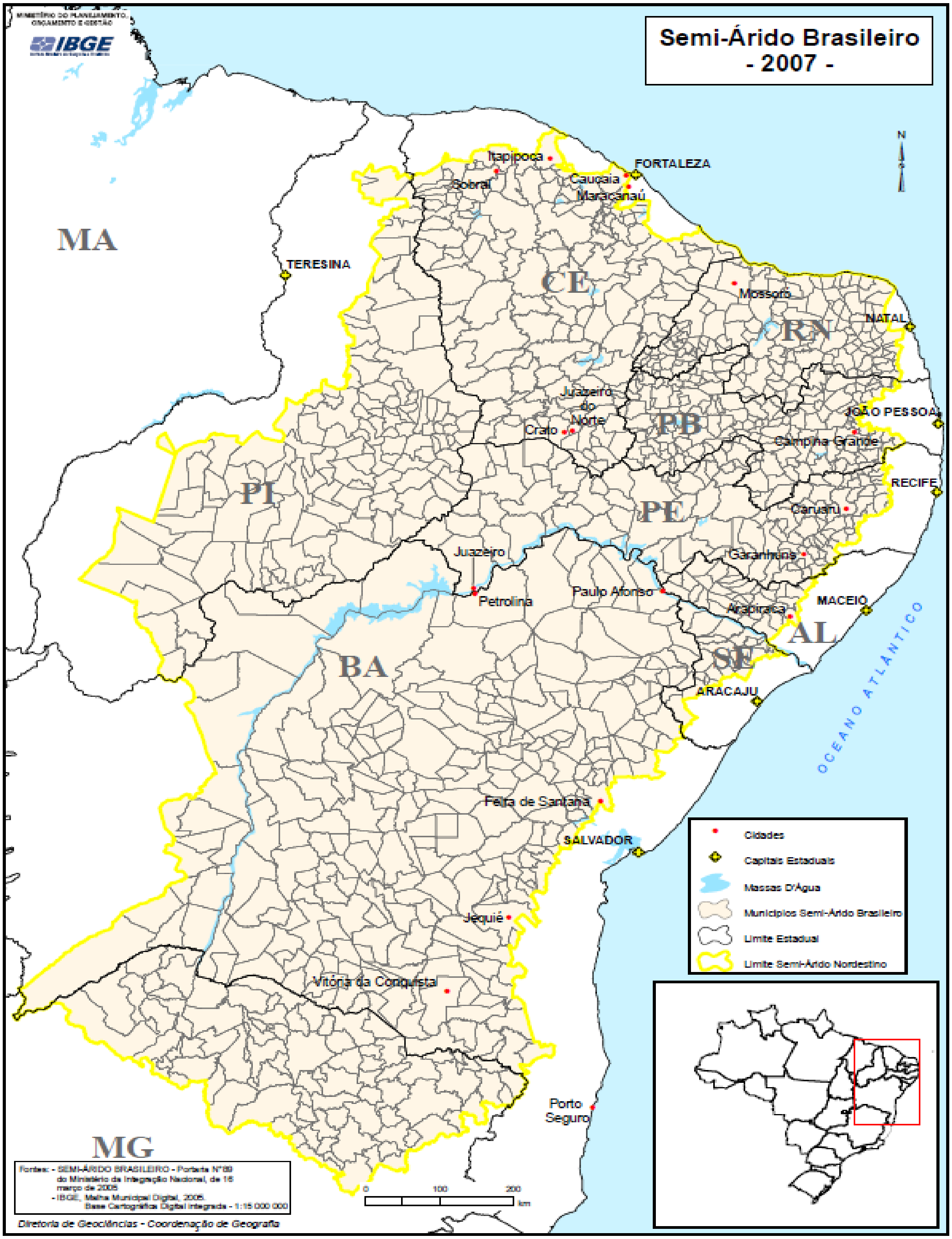


Figure 10: Map of the Semi-Arid portion of the Northeast Macroregion

There are at least 671 representative farms in every size and class, with the exception of family farms between 100ha and 500ha in the South. There are only 51 observations of this type, which generated imprecise and unstable econometric estimates. For this reason we exclude these large family farms in the South from the simulations. In the Semi-Arid region, we lost 20 percent of the representative farms due to aggregation and confidentiality, but only 7 percent of the individual farms. In the South we lost 21 percent of the representative farms, but only 3.4 percent of all farms.

Table 13: Representative Farms by Type and Size: Semi-Arid and South

	Semi-Arid			South		
	Representative farms	Farms (1000s)	Farms / Representative farms	Representative farms	Farms (1000s)	Farms / Representative farms
Total	7,144	1,481	207	6,821	926	136
Family (F)	3,796	1,330	350	3,282	790	241
0-5	964	705	732	1,054	190	180
5-20	988	346	350	1,073	380	354
20-100	1,027	247	240	1,104	218	198
100-500	817	32	39	51	1.9	36
Non-Family (NF)	3,348	151	45	3,539	136	38
0-5	671	54	81	675	19	28
5-20	814	32	39	924	29	31
20-100	922	34	37	1,023	39	38
100-500	941	31	33	917	49	54

Source: Agricultural Census-IBGE (2006)

4. What the Data Reveal: Empirical Results

4.1 Productivity, Farm Size and Farm Type: Descriptive Statistics

Tables 14 and 15 provide descriptive information on the data used in the model estimation. We caution that the data presented here have been filtered and, in some cases, defined differently than the variables discussed in previous sections. Thus, there is no reason for the data to match exactly. The data presented here are in levels and are based on regional averages. For this reason, they also differ from some of the results below which are based on logs and reflect the distributions of the variables, not just their means.

Consistent with the findings of Table 9, on average family farms have higher land productivity in both regions. In the Semi-Arid region land productivity of non-family farms is about 80 percent of the level of family farms, and in the South it is only about half. But as we argued in Section 2 of this Report, this is largely a result of differences in size. In both regions, family and non-family farms that have 0-5ha have roughly the same levels of productivity. In the Semi-Arid region, non-family farms above this size have between 50 percent and 140 percent higher land productivity. In the South, the difference only varies between 9 percent and 16 percent. But in all cases, non-family farms have higher land productivity.

In both regions, when we control for farm size, non-family farms use much more purchased inputs and capital per hectare than family farms.[11] In the Semi-Arid region, the differences are never less than 45 percent, and are often above 100 percent. Non-family farms between 20ha and 100ha, for example, use more than double the purchased inputs and capital per hectare than do family farms. In the South, the differences are less pronounced. Non-family farms used 25 percent to 30 percent more purchased inputs per hectare than family farms in each size class, and they used between 20 percent and 66 percent more capital per hectare. Differences in the use of family labor per hectare, in contrast, are quite small for farms under 20ha, but non-family farms with 20ha to 100ha use around 30 percent to 40 percent less family labor in both regions.

These differences in the intensity of input usage per hectare contribute to differences in land productivity across farms sizes and classes. Although non-family farms with 0-5ha of land used

[11] In the case of purchased inputs, it is important to recall that this includes hired labor.

much more purchased inputs and capital per hectare in both regions, and similar amounts of labor, land productivity was almost identical between these non-family and family farms. This might be suggestive of an efficiency advantage on family farms of this size. For the larger farms, between 5ha and 100ha, the more intensive use of purchased inputs and capital is sufficient to produce higher land productivity for non-family farms in spite of less family labor being used. This is especially true in the Semi-Arid.

The final column of Tables 14 and 15 shows profit per unit of family labor, calculated as the value of output minus the value of purchased inputs per unit of family labor of each size and class. This measure takes into account productivity (defined as value of output per hectare), farm size, variable costs, and the number of full-time equivalent family laborers of each farm size and type. It is important to emphasize that these are averages for each size and type, and the share that are poor depends on the distributions within each one. Nonetheless, the averages are suggestive. The annual poverty line that corresponds to one-half of a minimum wage per capita is about R$1,475.

The situation is quite different in the South, where both family and non-family farms spend much more on purchased inputs, have more capital, and produce at higher levels of productivity. On average, family farms in the 0-5ha class spent three times as much on inputs as similar farms in the Semi-Arid region, had four times as much capital, and generated nearly twice the value per hectare–all with less labor involved. Thus, short-run profits were more than R$600 higher per full-time family member, and approached 90 percent of the poverty line. Non-family farms with 0-5ha in the South only produced R$900 per equivalent family member, but as we saw in Table 8, a significant share of them earn considerable off-farm income. On average, family and non-family farms in the South in the larger size classes all generated sufficient income per full-time family member to earn multiples of the poverty line. Family members on family farms with 5-20ha earned 2.7 times the poverty line, and on family farms with 20-100ha this rose to 4.3 times the poverty line. Clearly, as the poverty data in Section 2 showed, agricultural income alone can lift most farms in the South out of poverty.

Table 14: Descriptive Statistics for the Semi-Arid

	Output	Purchased Inputs	Capita	Family labor	Profit
		(R$ per hectare)			(R$ per unit of family labor)
Family (F)	403	146	877	0.26	999
0-5	1,437	416	2,492	1.46	700
5-20	608	210	1,194	0.35	1,143
20-100	260	110	656	0.11	1,347
100-500	127	59	447	0.04	1,946
Non-Family (NF)	323	145	968	0.05	3,311
0-5	1,462	604	3,977	1.38	624
5-20	1,233	364	2,840	0.30	2,926
20-100	629	225	1,575	0.08	5,145
100-500	194	109	696	0.02	5,561
Ratio NF/F					
Total	0.80	0.99	1.10	0.21	3.32
0-5	1.02	1.45	1.60	0.94	0.89
5-20	2.03	1.73	2.38	0.85	2.56
20-100	2.42	2.04	2.40	0.71	3.82
100-500	1.53	1.86	1.56	0.44	2.86

Source: Agricultural Census-IBGE (2006)

Table 15: Descriptive Statistics for the South

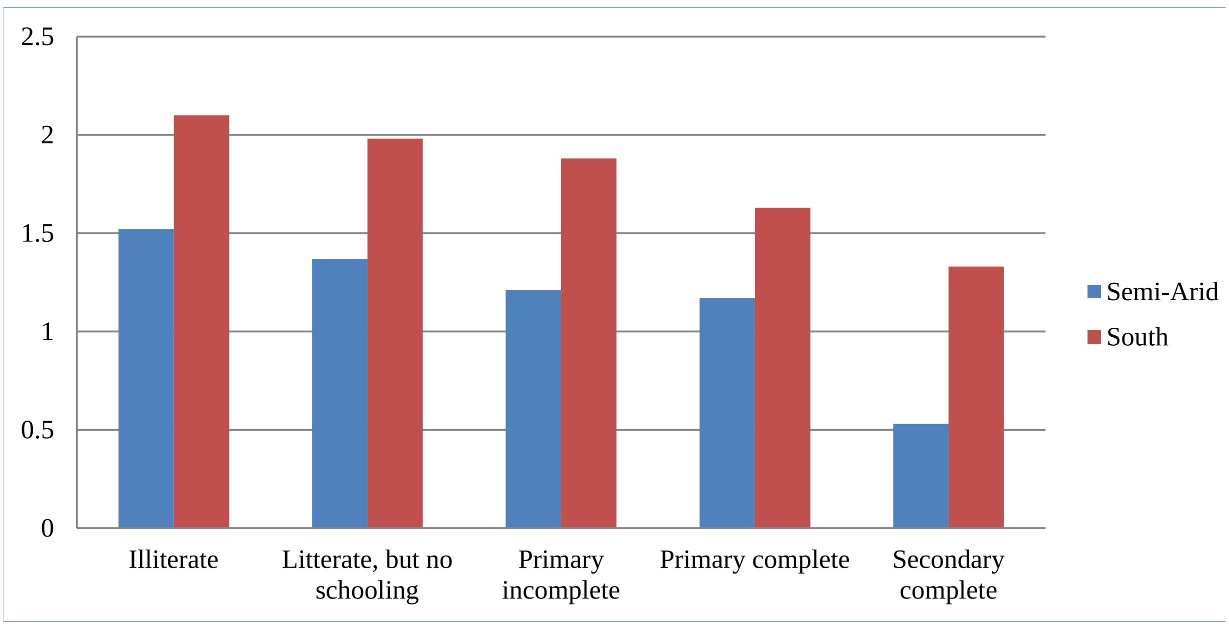
	Output	Purchased Inputs	Capital	Family labor	Profit
		(R$ per hectare)			(R$ per unit of family labor)
Family (F)	1,594	776	4,436	0.20	4,075
0-5	2,733	1,359	10,722	1.07	1,287
5-20	2,039	909	5,386	0.28	4,025
20-100	1,296	670	3,543	0.10	6,326
100-500	-	-	-	-	-
Non-Family (NF)	818	563	2,562	0.03	8,326
0-5	2,806	1,750	16,245	1.17	900
5-20	2,365	1,152	8,954	0.27	4,431
20-100	1,419	843	4,256	0.06	10,070
100-500	650	487	1,997	0.01	11,807
Ratio NF/F					
Total	0.51	0.73	0.58	0.15	2.04
0-5	1.03	1.29	1.52	1.10	0.70
5-20	1.16	1.27	1.66	0.97	1.10
20-100	1.09	1.26	1.20	0.58	1.59
100-500	-	-	-	-	-

Source: Agricultural Census-IBGE (2006)

4.2 Estimating the Model

The probit used in the first stage of the Heckman procedure worked quite well for distinguishing between family and non-family farms. In the Semi-Arid region, the model correctly classifies family and non-family farms 85 percent of the time, doing slightly better for non-family farms. In the South, the model correctly classifies the farms 84 percent of the time. Many of the correlations discussed in Section 2 of this Report reappear as important correlates of selection into family farming. For example, Figure 11 shows a strong negative relationship between

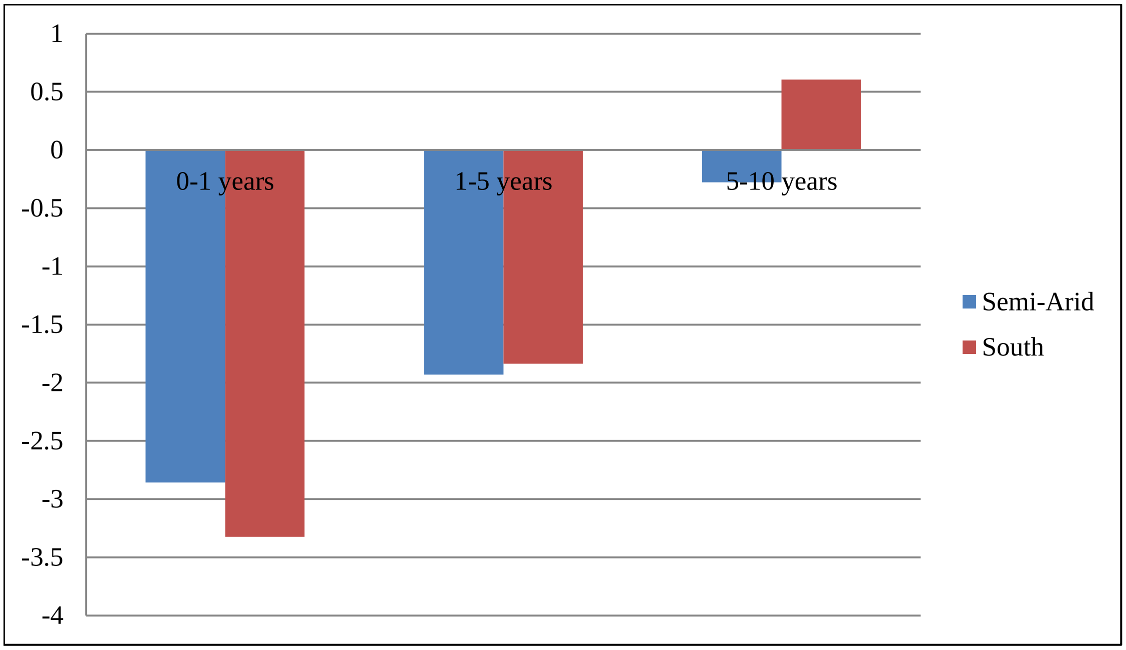
Figure 11: Schooling and the Probability of Being a Family Farmer



Source: Agricultural Census-IBGE (2006)

schooling and the probability of family farming. Figure 11 shows the marginal effects calculated from a model in which the excluded class is family farms with a higher education, and the right hand-side variable is the percentage of farms within each representative farm that has a particular level of schooling. Thus, relative to producers with a higher education, a one percentage point increase in the share of farmers who are illiterate or who have no formal schooling is associated with an increase of around 1.5 percentage points in the probability of becoming family farmers in the Semi-Arid, and an increase of around two percentage points in the South.

Figure 12: Experience and the Probability of Being a Family Farmer



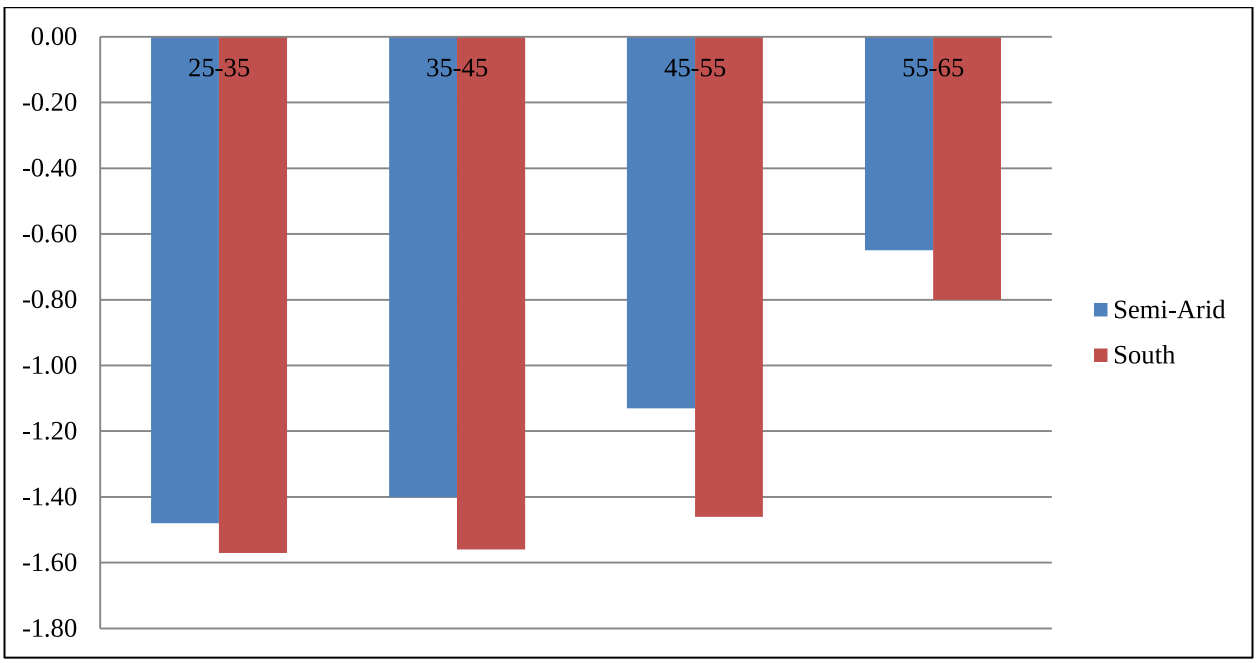
Source: Agricultural Census-IBGE (2006)

Figure 12 shows that relative to farmers with 10 or more years of experience managing the current farm, farmers who have taken charge of their farms more recently are more likely to be non-family in both regions. A one percentage point increase in the share of farmers who have been in charge of the current farm for less than one year is associated with a decrease of between 2.75 and 3.40 percentage points in the probability of being a family farmer for the Semi-Aird Northeast and the South, respectively. Farmers with 1-5 years of experience—rather than 10—are also less likely to be family farmers, but the marginal effect is smaller in both regions.

Figure 13 shows a strong positive relationship between age and the probability of being a family farmer. Younger farmers are significantly less likely to be family farmers in both regions. Taken as a whole the evidence from these Figures 11-13 suggests a changing of the guard, with older, less educated farmers continuing on as family farmers, and younger, more educated—and perhaps more entrepreneurial—farmers entering the ranks as non-family farmers.

The coefficient on the inverse Mills ratio is consistent with the above view. When the inverse Mills ratio is calculated from the first stage and included in the second stage estimation of the translog production function, it is significant at least at the 1 percent level in both regions and for both the family and non-family models. Interestingly, the coefficient (λ) is positive in both non-family models and negative in both family production functions. This suggests that there are unobservables that favor both (i) selection into the non-family farm group and (ii) higher levels of land productivity. The reverse is true for family farms. The unobservables are associated with an increased likelihood of selection into the family farm group and with lower levels of land productivity. While we did control for a rich set of human capital variables, including age, experience, and education, it is possible that there are unobserved variables, such as risk taking and entrepreneurship, that lead some farmers to hire labor in or out—and thus become a non-family farm—and at the same time to achieve a higher than predicted level of productivity.

Figure 13: Age and the Probability of Being a Family Farmer



Source: Agricultural Census-IBGE (2006)

4.3 Elasticity of Output per Hectare with Respect to Inputs per Hectare

Table 16 shows output elasticities with respect to inputs in the Semi-Arid and South macroregions. Because elasticities are variable with a translog, and depend on the levels of the inputs, we calculate the elasticities at the means of all variables. As expected, the family labor elasticity is close to zero for most sizes in both regions. This suggests that family labor is not constrained, and might even be oversupplied in some cases due to an abundance of family labor with few alternative employment opportunities.

Table 16: Elasticities of Output per Hectare with Respect to Inputs per Hectare

	Semi-Arid			South		
	Family labor	Capital	Purchased Inputs	Family labor	Capital	Purchased Inputs
Family (F)						
0-5	0.15	0.32	0.94	0.01	0.44	0.01
5-20	0.01	0.49	0.85	-0.02	0.59	0.05
20-100	0.00	0.57	0.69	-0.05	0.61	0.20
100-500	0.02	0.59	0.66	-	-	-
Non-Family (NF)						
0-5	0.11	0.49	1.25	0.02	0.44	0.18
5-20	0.08	0.66	1.15	-0.04	0.69	0.23
20-100	0.03	0.78	0.93	-0.05	0.76	0.39
100-500	0.15	0.77	1.00	0.10	0.48	0.67

Source: Agricultural Census-IBGE (2006)

In the Semi-Arid Northeast, capital and input elasticities are quite high, reflecting the low levels at which both of these inputs are used in the region, and the potential gains at the margin if these factors were expanded. In the Semi-Arid Northeast, purchased input elasticities for family farms (i) fall with farm size; and (ii) are more than those for capital at all farm sizes. In the South, in contrast, purchased input elasticities (i) rise with farm size; and (ii) are uniformly less than those for capital at all farm sizes. Finally, for non-family farms in both regions, both capital and purchase input elasticities exceed those for family farms across all farm sizes.

These findings provide clues about which inputs could have the largest impact on output per hectare for each type and size of farm. In the Semi-Arid Northeast, the smallest farms, whether family or non-family, would appear to be able to benefit substantially from expanding their use of purchased inputs. The impacts of increasing the use of capital are all considerable. The pattern is somewhat different in the South macroregion. The capital elasticities are substantially higher than the purchased input elasticities in all cases but one.

4.4 Decomposing Land Productivity: Factor Intensity, Non-factor Variables and Technology

Table 17 and 18 decompose the differences in land productivity ($y_{nf} - y_f$) into differences in: (i) the intensity in the use of factors of production (dFat); (ii) non-factors such as climate, soils, human capital, credit and technical assistance (dNfat); and (iii) technology as described by the coefficients in the production function (dTec). The difference in total factor productivity (dTFP) between two groups equals the sum of dTec and dNfat. The decompositions are presented for two cases: (i) family vs. non-family, conditional on size, and (ii) neighboring size classes, k vs. k+1, conditional on family type. Table 17 and 18 show the decompositions for the Semi-Arid Northeast and the South, respectively. We first discuss the family vs. non-family decomposition. The first column in Table 17 shows the difference in average log land productivity ($y_{nf} - y_f$) between non-family and family farms. Consistent with what was presented descriptively above, in each size class non-family farms have higher average log land productivity than family farms.

Table 17: Decomposition of Differences in Land Productivity in the Semi-Arid

	Δ Family Type Conditional on Size					Δ Size Conditional on Type				
	y_{nf} - y_f	dFat	dTec	dNfat	dTFP	y_{k+1} - y_k	dFat	dTec	dNfat	dTFP
	difference in logs					difference in logs				
Family (F)										
0-5	0.06	0.52	-0.67	0.22	-0.45	-0.71	-0.96	0.34	-0.09	0.25
5-20	0.25	0.70	-0.67	0.24	-0.44	-0.72	-0.72	0.09	-0.09	0.00
20-100	0.36	0.74	-0.61	0.25	-0.36	-0.78	-0.69	0.24	-0.32	-0.09
100-500	0.03	0.46	-0.63	0.18	-0.45	-	-	-	-	-
Non-Family (NF)										
0-5	-					-0.52	-1.16	0.69	-0.06	0.63
5-20		-				-0.61	-0.99	0.57	-0.18	0.39
20-100			-			-1.10	-1.35	0.51	-0.27	0.24
100-500				-	-	-	-	-	-	-

Source: Agricultural Census-IBGE (2006)

The difference in logs ranges from .03 to .36, with the largest difference found in the 5-20 and 20-100 size classes. These differences in land productivity are largely explained by higher factor intensities for non-family farms (dFat). Factor intensities are between 46 percent and 74 percent larger for non-family farms in all size classes. But since factor intensity increases much more than land productivity, something must be offsetting this increase. This is reflected in a TFP advantage for family farms (dTFP) on the order of 36 percent to 45 percent. And when we decompose the TFP advantage of family farms, we see that it comes from technology (dTec) rather than from non-factors (dNfat). Thus, family farms in the Semi-Arid use inputs much less intensively, but do a relatively better job at converting these into output per hectare.

A similar result can be seen when we let each individual representative farm choose the optimal technology (family vs. non-family) conditional on farm size. Here, the simulations hold inputs and non-factors constant and examine which vector of technology coefficients maximizes output. More than 80 percent of the representative non-family farms choose the family farm technology, and this leads to a considerable increase in yields. Between 82 percent and 96 percent of the representative family farms, in contrast, choose to stick with the technology that they currently use. This suggests a lack of inefficiency for this group.

In the South macroregion (Table 18), we see much smaller differences in log land productivity between family and non-family farms ($y_{nf} - y_f$). Ignoring the 100-500ha class because of the small number of family farm observations, the log differences in land productivity are 0.09 for the 5-20ha class and close to zero in the other cases. As in the Semi-Arid, the non-family farms use inputs more intensively, but don't manage to convert this into a proportional increase in land productivity. Non-family farms use inputs between 12 percent and 27 percent more intensively, but their TFP is estimated to be 15 percent to 19 percent lower. Like in the Semi-Arid, when we decompose TFP the family farm advantage comes from the technology, not the non-factor inputs. Thus, non-family farms have an advantage in terms of inputs and non-factor variables, but family farms use their scarce inputs efficiently.

The results are once again confirmed by the second experiment. Holding inputs and non-factors constant, more than 77 percent of non-family farms in each size class could increase land productivity by switching to the technology of the family farms. Less than 11 percent of family farms can increase productivity by switching technologies.

The decomposition results are also used to explore the reasons for the differences in land productivity between farms of the same type that are of neighboring size classes. Thus, we compare each size (k) to the size immediately above it ($k+1$). Because of the well-known inverse relationship between farm size and productivity, we find, as expected, that larger farms systematically have lower land productivity. Table 17 shows that yields (y_{k+1} - y_k) fall between 52 percent and 110 percent with each increase in farm size class in the Semi-Arid, and in the South (Table 18) the reductions range from 22 percent to 72 percent.

Table 18: Decomposition of Land Productivity in the South

	Δ Family Type Conditional on Size					Δ Size Conditional on Type				
	y_{nf} - y_f	dFat	dTec	dNfat	dTFP	y_{nf} - y_f	dFat	dTec	dNfat	dTFP
Family (F)										
0-5	0.00	0.12	-0.54	0.40	-0.15	-0.31	-0.37	-0.02	0.08	0.06
5-20	0.09	0.27	-0.51	0.34	-0.18	-0.35	-0.19	-0.07	-0.09	-0.16
20-100	-0.01	0.17	-0.44	0.25	-0.19	-	-	-	-	-
100-500	-	-	-	-	-	-	-	-	-	-
Non-Family (NF)										
0-5	-					-0.22	-0.41	0.24	-0.05	0.19
5-20		-				-0.45	-0.45	0.15	-0.15	0.00
20-100			-			-0.72	-0.69	0.33	-0.36	-0.03
100-500				-	-	-	-	-	-	-

Source: Agricultural Census-IBGE (2006)

Note: The simulation from NF to F equals -1 times the simulation from F to NF.

The explanations for the inverse relationship, however, differ across regions. In the Semi-Arid, the input intensity falls much more rapidly than the yields in four of six cases. In these four cases, each successively larger farm size class compensates for the decline in input intensity through higher levels of technology that are reflected in higher TFP. The TFP advantage of class k+1 over class k varies from 25 percent to 63 percent in these cases. In the South there is only one case—non-family with 0-5ha—where input intensity falls much more rapidly than land productivity. The other comparisons in the South show declines in yields that are comparable to the declines in inputs. In the South, differences in input intensities almost fully explain the declines in yields. Thus, there are much smaller differences in technology and TFP across neighboring size classes. The largest difference in TFP in the South is 19 percent, whereas it surpasses 60 percent in the Semi-Arid. For family farms in the South, the technology actually appears to decline slightly as we move from one size class to the next.

The results are once again confirmed by the simulations that allow each representative farm to choose among technologies of the same farm type (family vs. non-family). Because technologies in the Semi-Arid improve monotonically with farm size, all family and non-family farms would prefer to use the technology of the largest size class in their respective group. This is also true for non-family farms in the South. However, the results are somewhat different for family farms in the South. About a quarter of the smallest farms, and over one third of the farms in the 5-20ha class, maximize output with the technology that their group currently uses. This is consistent with the results of the decomposition above. Family farms in the 20-100ha class do not appear to be using a superior technology to farms in the 0-5ha class.

5. Dissecting the Differences Among Family Farms: Extensions to the Base Model

The previous sections of this Report explore differences between family and non-family farms. In this section, we restrict the analysis to family farms because the overwhelming majority of poor farmers in Brazil are of this type. Within the family farm category, the model is extended to explore differences between farms across the following dimensions: (i) use of credit; (ii) uptake of technical assistance; (iii) specialization of production; and (ii) use of irrigation.[12] For each classification, we present information on four fronts: (i) poverty; (ii) use of inputs; (iii) factors associated with the selection into each group; and (iv) productivity. The model that was defined in Section 3 of this Report, and estimated for family vs. non-family farmers by farm size, is applied to the classifications studied here.

5.1 Credit, Technical Assistance, Specialization and Irrigation

Table 19 shows the number of farms and representative farms by type in the Semi-Arid and South of Brazil. As reported in Table 13, family farms account for 90 percent of the farms in the Semi-Arid and 85 percent in the South. Of the 1.33 million family farms in the Semi-Arid, only 15 percent report having used credit in 2006, 8 percent say that they used technical assistance, 55 percent were classified as "specialized" in the sense of having at least 65 percent of the value of their output in a single product, and 6 percent used irrigation. The situation in the South was quite different. Of the 790,000 family farmers, 38 percent used credit, 48 percent used technical assistance, and 56 percent were specialized.[13] Thus, credit was more than twice as common in the South, and the use of technical assistance was six times more likely. As we will see, these two factors are correlated with higher productivity and lower poverty levels in both regions.

---

[12] For irrigation, due to data anomalies, only a descriptive presentation is made.

[13] Because water scarcity is a much more pressing issue in the Semi-Arid, we focused on this region and did not attempt to estimate the irrigation model in the South.

For those producers who did not use credit, the 2006 Agricultural Census had a novel question asking about why they did not borrow any money. Although we do not have this variable disaggregated by type, we do know that in the Semi-Arid only 16 percent of those who did not borrow said that it was for one of the following reasons: (i) did not have collateral (2 percent); (ii) did not know how to get a loan (one percent); (iii) bureaucracy (8 percent); or (iv) failure to pay a previous loan (5 percent). The most important reasons were "did not need it" (39 percent) and fear of taking on debt (27 percent). Another 17 percent cited "other reasons." Thus, it seems that there is scope for expanding credit accompanied with some sort of insurance that would address the fears of over one-quarter of the producers in the Semi-Arid.

It is also worth investigating what lies behind the response "did not need it." It is possible that some producers can self-finance, but many family farms operate at close to a subsistence level, and do not even contemplate expansion, modernization, or other strategies that might boost their productivity and income. Again, if coupled with insurance, there is likely scope for expanding credit within this group of producers as well. In the South, 72 percent of all producers (family and non-family) who did not borrow said that they did not need it. Given the much higher rate of credit access in the South (38 percent of family farmers), it is more plausible to take this response at face value. Only 12 percent cited fear of taking on debt in the South.

Table 19: Number of Farms and Representative Farms by Type (Semi-Arid and South)

	Semi-Arid			South		
	Representative Farms	Farms (1000s)	Share of Farms	Representative farms	Farms (1000s)	Share of Farms
Total	7,144	1,481		6,821	926	
Family (F)	3,796	1,330	0.90	3,282	790	0.85
Non-Family (NF)	3,348	151	0.10	3,539	136	0.15
Family Farms:						
Credit (C)	3,001	193	0.15	2,880	301	0.38
No Credit (NC)	3,752	1,136	0.85	3,263	488	0.62
Tech. Assistance (TA)	3,306	103	0.08	5,641	375	0.48
No Tech. Assistance (NTA)	3,730	1,224	0.92	3,161	414	0.52
Specialized (S)	6,015	654	0.55	5,200	410	0.56
Not Specialized (NS)	5,327	540	0.45	4,980	323	0.44
Irrigation (I)	1,833	78	0.06	---	---	---
No Irrigation (NI)	3,753	1,251	0.94	---	---	---

Source: Agricultural Census-IBGE (2006)

5.2 Poverty, Inputs and Non-Factor Variables among Family Farms

Tables 20 and 21 show poverty rates for family farms by size in the Semi-Arid Northeast and South, respectively. Table 20 shows that 73 percent of family farms are poor in the Semi-Arid when only agricultural income is considered. Furthermore, poverty falls from 82 percent down to 51 percent as farm size increases from the 0-5ha class up to the 100-500ha class. When family farms are disaggregated according to their use of credit, the difference is not particularly large: overall, those that used credit (only 15 percent of these family farms) had a poverty rate that was five percentage points lower. It is important to emphasize that this cannot be interpreted as the causal impact of credit. It likely reflects a combination of a causal impact and self-selection. Those that chose to use credit might be different in observable and unobservable ways, might have had a lower poverty rate anyway, or might make better use of credit which is part of the reason why they chose to use it. The gap in poverty rates vis-à-vis use of credit is fairly constant across farm sizes, ranging from 4 to 6 percentage points.

Strikingly, as Table 19 shows, only 8 percent of family farms in the Semi-Arid Northeast chose to use technical assistance; this is about one-sixth of the same figure for family farms in the South. Family farms in the Semi-Arid that chose to use technical assistance had a poverty rate that was 15 percentage points lower than those who did not use technical assistance (60 percent vs. 75 percent), and the poverty rate for farms that used both credit and technical assistance was 20 percentage points lower than the poverty rate of farms that used neither (56 percent vs. 76 percent). Thus, poverty falls from 68 percent to 56 percent when we compare farms that used credit to those that used both credit and technical assistance. Comparing technical assistance alone to technical assistance and credit, poverty falls from 60 percent down to 56 percent. Choosing to use technical assistance appears to be more important than choosing to use credit.

Specialized family farms i.e., with at least 65 percent of their value of output in a single crop, have somewhat lower poverty than diversified farms. The difference is nine percentage points. Only 57 percent of the farms that used irrigation are poor. Poverty is lower by 18 percentage points for farms using irrigation. Yet and still, only 6 percent of family farms in the Semi-Arid used irrigation. These findings, while not causal, are highly suggestive of the importance of irrigation in the Semi-Arid Northeast in closing the poverty gap.

Table 20: Poverty Rates (%) for Family Farms by Farm Size and Non-Factor Variables (Semi-Arid Northeast)

Type	% Poverty by type	% Poverty by Farm Size				% of farms by type
		0-5	5-20	20-100	100-500	
Family	73	82	69	59	51	90
Credit						
Yes	68	78	63	55	45	15
No	74	83	70	60	52	85
Technical Assistance						
Yes	60	70	59	51	41	8
No	75	83	70	60	52	92
Credit and TA						
Both	56	67	55	50	36	2
Neither	76	83	71	61	53	79
Specialized[2]						
Yes	66	78	60	49	40	54
No	75	84	71	61	50	46
Irrigation						
Yes	57	65	53	44	37	6
No	75	83	70	60	52	94

Source: 2006 Agricultural Census
Notes:
1. This table uses the short run poverty measure with adult equivalent full time workers.
2. 65% or more of value of output in principal product

Relative to the Semi-Arid Northeast, Table 21 for the South shows much larger differences in poverty rates overall and for those of credit and technical assistance, a smaller difference with respect to irrigation, and essentially no gap in the case of specialization. Poverty among these family farms, at 40 percent, is 33 percentage points less than family farms in the Semi-Arid Northeast. Across farm sizes in the South, poverty drops by 27 percentage points between 0-5 ha and 5-20ha; this decrease is slightly more than double the decrease between these same farm sizes in the Semi-Arid Northeast.

Interestingly, family farmers in the South that use credit or access technical assistance have one-half the poverty of those farms without these services, and farms that used both credit and technical assistance have one-third the poverty of those that use neither (59 percent vs. 20 percent). Not only are credit and technical assistance strongly correlated with lower poverty, but farmers in the South use these services at a much higher rate: 26 percent of family farmers in the South used both services, while only 2 percent in the Semi-Arid Northeast did so.

Only 8 percent use technical assistance in the Semi-Arid Northeast, while 48 percent did in the South. Specialization in the South was associated with lower poverty for smaller farms and slightly higher poverty for larger farms. It is possible that this is a reflection of small farms that are integrated into chicken, hog and other agroindustry in the South. Perhaps because scarcity of water is not as severe a constraint in the South vis-à-vis the Semi-Arid Northeast, the difference in poverty between those who have and do not have irrigation in the South was not as large as it was in the Semi-Arid. In what follows, we do not continue to investigate irrigation in the South.

Table 21: Poverty Rate (%) for Family Farms by Farm Size and Non-Factor Variables (South)

Type	% Poverty by type	Farm Size			% of farms by type
		0-5	5-20	20-100	
Family	40	62	35	30	85
Credit					
Yes	24	45	23	20	38
No	50	66	44	40	62
Technical Assistance					
Yes	26	43	25	22	48
No	53	70	46	42	52
Credit and TA					
Both	20	33	20	18	26
Neither	59	72	52	49	40
Specialized[2]					
Yes	36	55	31	27	56
No	35	64	31	23	44
Irrigation					
Yes	29	49	25	19	5
No	41	63	36	31	95

Source: 2006 Agricultural Census
Notes:
1. This table uses the short run poverty measure with adult equivalent full time workers.
2. 65% or more of value of output in principal product

5.3 Land Productivity, Use of Inputs, and Profit per unit of Family Labor

Tables 22-27 are comparable to Tables 14-15 for the Semi-Arid Northeast and South, but rather than comparing farm type and farm size in the aggregate, they draw these comparisons for only family farms by disaggregating across credit, technical assistance, and specialization. For the Semi-Arid Northeast, disaggregation is also in regard to irrigation.

The comparison of family farms with respect to the use of credit reveals some interesting contrasts. In the Semi-Arid Northeast and across farm sizes, family farmers that use credit generate yields (i.e., land productivity as measured in R$ per hectare) that are between 30 percent and 43 percent higher as a result of greater use of capital, inputs, and family labor. On average, these variables are 23 percent, 9 percent and 16 percent higher, respectively, than for family farms that do not use credit. As a result, on average, short-run profit per family member is 34 percent higher for farms using credit.

In the South, land productivity is 60 percent higher for those farms that use credit, and most of the yield gap comes from greater use of inputs. Input usage is 40 percent higher for farms using credit. As a result of these differences, farms in the 0-5ha class that use credit, for example, generate more than double the profit per unit of family labor. By contrast, credit use by family farms in the Semi-Arid Northeast for this same farm size is associated with only a 35 percent increase in profit per unit of family labor. As section 1.5 of this Report noted, investment contracts under PRONAF – one indication of the magnitude of credit use – differ between the South and the Northeast (which includes the Semi-Arid) by a factor of twenty.

The yield gap for family farms in the Semi-Arid Northeast, in regard to the use of technical assistance is larger than that observed in the case of the use of credit (Table 24). Farms with technical assistance have yields that are between 70-100 percent higher across size classes, and the use of capital appears to be the most important variable in achieving this. Farms between 0 and 20 ha that use technical assistance deploy nearly twice as much capital, around 25 percent more inputs, and 5-15 percent more family labor. The combined result is profit per unit of family labor that is more than double that of farms of this size that do not use technical assistance. In the South, the differences in yields and short-run profit per unit of family labor are even larger (Table 25). In this region, the differences in purchased inputs are of the same magnitude as the differences in capital.

Table 22: Family Farms in the Semi-Arid: Credit vs. No Credit

	Output	Purchased Inputs	Capital	Family labor	Profit
		(R$ per hectare)			(R$ per unit of family labor)
Credit (C)					
All Family Farms	536	147	1,054	0.30	1,317
0-5	1,890	487	3,115	1.61	873
5-20	822	203	1,437	0.42	1,481
20-100	325	101	759	0.13	1,770
100-500	164	55	473	0.04	3,044
No Credit (NC)					
All Family Farms	384	135	855	0.25	982
0-5	1,390	446	2,435	1.46	647
5-20	574	188	1,160	0.34	1,139
20-100	249	95	641	0.11	1,415
100-500	121	53	442	0.03	1,951
Ratio C/NC					
All Family Farms	1.40	1.09	1.23	1.16	1.34
0-5	1.36	1.09	1.28	1.10	1.35
5-20	1.43	1.08	1.24	1.24	1.30
20-100	1.30	1.06	1.18	1.16	1.25
100-500	1.35	1.02	1.07	1.03	1.56

Source: Agricultural Census-IBGE (2006)

Table 23: Family Farms in the South: Credit vs. No Credit

	Output	Purchased Inputs	Capital	Family labor	Profit
		(R$ per hectare)			(R$ per unit of family labor)
Credit (C)					
All Family Farms	2,005	865	4,753	0.19	6,082
0-5	4,063	1,907	10,226	1.01	2,139
5-20	2,548	986	5,717	0.29	5,427
20-100	1,673	776	4,115	0.11	8,133
No Credit (NC)					
All Family Farms	1,251	618	4,189	0.21	2,981
0-5	2,375	1,328	10,918	1.09	964
5-20	1,627	698	5,131	0.28	3,371
20-100	944	509	3,017	0.09	4,930
Ratio C/NC					
All Family Farms	1.60	1.40	1.13	0.88	2.04
0-5	1.71	1.44	0.94	0.93	2.22
5-20	1.57	1.41	1.11	1.05	1.61
20-100	1.77	1.53	1.36	1.25	1.65

Source**:** Agricultural Census-IBGE (2006)

Table 24: Family Farms in the Semi-Arid: Technical Assistance vs. No Technical Assistance

	Output	Purchased Inputs	Capital	Family labor	Profit
		(R$ per hectare)			(R$ per unit of family labor)
Technical Assistance (TA)					
All Family Farms	621	156	1,288	0.22	2,082
0-5	2,841	570	4,497	1.54	1,476
5-20	1,127	242	2,153	0.40	2,219
20-100	410	123	963	0.12	2,440
100-500	206	67	625	0.04	3,596
No Tech. Assist. (NTA)					
All Family Farms	382	138	837	0.27	919
0-5	1,351	449	2,367	1.46	619
5-20	559	192	1,102	0.35	1,060
20-100	242	96	619	0.11	1,324
100-500	116	51	420	0.03	1,878
Ratio TA/NTA					
All Family Farms	1.63	1.12	1.54	0.84	2.27
0-5	2.10	1.27	1.90	1.05	2.38
5-20	2.02	1.26	1.95	1.15	2.09
20-100	1.70	1.29	1.55	1.06	1.84
100-500	1.78	1.31	1.49	1.13	1.91

Source: Agricultural Census-IBGE (2006)

Table 25: Family Farms in the South: Technical Assistance vs. No Technical Assistance

	Output	Purchased Inputs	Capital	Family labor	Profit (R$)
		(R$ per hectare)			(R$ per unit of family labor)
Technical Assistance (TA)					
All Family Farms	2,082	808	5,187	0.18	6,937
0-5	4,464	1,958	13,370	0.99	2,524
5-20	2,715	957	6,527	0.28	6,188
20-100	1,702	698	4,264	0.10	9,574
No Tech. Assist. (NTA)					
All Family Farms	986	570	3,526	0.22	1,856
0-5	1,903	1,162	9,509	1.11	665
5-20	1,308	648	4,167	0.28	2,365
20-100	694	462	2,490	0.09	2,562
Ratio TA/NTA					
Total	2.11	1.42	1.47	0.82	3.74
0-5	2.35	1.69	1.41	0.89	3.80
5-20	2.08	1.48	1.57	1.02	2.62
20-100	2.45	1.51	1.71	1.16	3.74

Source: Agricultural Census-IBGE (2006)

Tables 26 and 27 are equally illuminating. Specialized farms use roughly the same amount of purchased inputs, capital and family labor as non-specialized farms, and this is true in both regions. The large differences in value of output per hectare appear to be related to the fact that they have *specialized in what are probably higher-value crops*. The ability to specialize, in turn, leads to profits per unit of family labor that are at least double in all size classes and in both regions, with the exception of the 20-100ha farms in the South where the differences is 86 percent. The question, then, is what leads—or permits—some producers to make this choice while others do not? Here, the significant differential in accessing technical assistance between family farms in the South and Semi-Arid Northeast (48 percent vs. 8 percent) may help explain the orientation toward higher-value crops. Furthermore, we hypothesize that it is related to entrepreneurship, willingness to bear risk, and access to institutions that help shield family farmers from risk. Examples of such institutions are cooperatives, contract farming, or integration into high value agroindustrial commodity chains.

Finally, as expected, irrigated family farms in the Semi-Arid are quite different than those without irrigation (Table 28). They tend to have at least twice as much capital, use around 50 percent more purchased inputs, and produce three times as much value per hectare. It is likely that they are specialized in higher value crops as well. As a result, average profit per unit of family labor is below the poverty line for all farm sizes when they don't use irrigation, and is around four times higher in each size class when they do use irrigation. Access to water is strongly associated with positive outcomes.

Table 26: Family Farms in the Semi-Arid: Specialized vs. Not Specialized

	Output	Purchased Inputs	Capital	Family labor	Profit
		(R$ per hectare)			(R$ per unit of family labor)
Specialized (SP)					
Total	564	126	955	0.25	1,754
0-5	2,058	429	2,824	1.50	1,084
5-20	879	176	1,323	0.34	2,087
20-100	362	89	710	0.11	2,584
100-500	176	52	476	0.03	3,659
Not Specialized (NSP)					
Total	336	125	930	0.31	671
0-5	1,188	420	2,536	1.65	464
5-20	475	171	1,249	0.42	729
20-100	211	83	689	0.13	972
100-500	105	47	452	0.04	1,447
Ratio SP/NSP					
Total	1.68	1.01	1.03	0.80	2.61
0-5	1.73	1.02	1.11	0.91	2.34
5-20	1.85	1.03	1.06	0.81	2.86
20-100	1.71	1.08	1.03	0.80	2.66
100-500	1.68	1.11	1.05	0.85	2.53

Source: Agricultural Census-IBGE (2006)

Table 27: Family Farms in the South: Specialized vs. Not Specialized

	Output	Purchased Inputs	Capital	Family labor	Profit
		(R$ per hectare)			(R$ per unit of family labor)
Specialized (SP)					
Total	2,060	663	4,712	0.21	6,731
0-5	3,747	1,406	11,375	1.08	2,175
5-20	2,728	747	5,945	0.28	6,971
20-100	1,581	565	3,557	0.10	10,296
Not Specialized (NSP)					
Total	1,316	648	4,440	0.21	3,226
0-5	1,839	1,225	10,531	1.14	540
5-20	1,526	711	5,061	0.29	2,779
20-100	1,188	587	3,806	0.11	5,529
Ratio SP/NSP					
Total	1.57	1.02	1.06	1.00	2.09
0-5	2.04	1.15	1.08	0.95	4.03
5-20	1.79	1.05	1.17	0.97	2.51
20-100	1.33	0.96	0.93	0.91	1.86

Source: Agricultural Census-IBGE (2006)

Table 28: Family Farms in the Semi-Arid: Irrigation vs. No Irrigation

	Output	Purchased Inputs	Capital	Family labor	Profit
		(R$ per hectare)			(R$ per unit of family labor)
Irrigation (I)					
Total	1,111	209	1,890	0.28	3,246
0-5	3,956	722	5,624	1.56	2,075
5-20	1,986	306	3,233	0.46	3,680
20-100	676	148	1,273	0.11	4,701
100-500	301	79	758	0.03	6,580
No Irrigation (NI)					
Total	358	143	813	0.26	834
0-5	1,286	465	2,306	1.46	563
5-20	523	200	1,069	0.34	945
20-100	233	100	616	0.11	1,194
100-500	113	56	420	0.04	1,623
Ratio I/NI					
Total	3.11	1.46	2.33	1.08	3.89
0-5	3.08	1.55	2.44	1.07	3.69
5-20	3.80	1.53	3.02	1.33	3.90
20-100	2.90	1.47	2.07	1.01	3.94
100-500	2.66	1.41	1.80	0.96	4.05

Source: Agricultural Census-IBGE (2006)

5.4 Selection into Types: The Role of Gender, Human Capital and Non-Factor Variables

The heterogeneity of family farmers was described in the previous sections by analyzing differences in poverty and input and output intensities, across farm sizes and types. These differences are the result of observed and unobserved factors that influence the choice of technology and efficiency. We now explore the marginal effects of several key variables on the probability of: (i) being a family farmer; (ii) choosing to use credit and technical assistance; or (iii) specializing in production. Because the marginal effects are estimated by including farm size and other variables, they control for differences in many dimensions and thus compare producers that are similar based on observables. The marginal effects can be thought of as conditional correlations; they should not be interpreted as a causal.

Table 29 shows the marginal effects of gender, schooling, experience and age on the probability of choosing to be a family farmer, to use credit or technical experience, or to specialize in production. Because the models are estimated with aggregate data, and the explanatory variables are measured as the share of producers within a particular category, the marginal effects should be interpreted as follows: in the case of illiterate producers and credit, for example, they show the impact on the probability of choosing to use credit (in relation to the excluded category of farmers—those with more than a secondary education) that results from a one percentage point increase in the share of illiterate producers. The use of categorical variables permits capturing non-linear relationships.

The first two columns of Table 29 reproduce the information that was presented in Figures 10 through 12. They show that, in both regions, the probability of being a family farmer *decreases* successively with increased schooling, and *increases* successively with increases in age and years in charge of the current farm. By reading across the rows of the tables, we can examine the relationship between a particular characteristic and the probability of choosing the alternative categories. Women in both regions, for example, are more likely to be family farmers, are less likely to use credit and technical assistance, and are less likely to specialize in production.

The marginal effects of years in charge of the current farm show that farmers with more than 10 years of experience are much more likely to use credit, to specialize in production, and (in the South) to use technical assistance, while farmers with 0-1 years of experience on the current farm are the least likely to make these choices. As the age of producers increases in both regions, the probability of using credit decreases and the probability of specializing increases. These trends are less linear in the Semi-Arid Northeast. Producers over 65 are the least likely to use technical assistance in both regions, although the probability falls fairly consistently with age in the South, yet remains relatively constant in the Semi-Arid.

The relationship between schooling and selection is more complex and non-linear, and appears to reflect the influence of the PRONAF program. Producers with some form of post-secondary education are the least likely to use credit and the most likely to use technical assistance. Curiously, together with farmers who have a complete secondary education, they are also the least likely to specialize in production. The probabilities of selection rise and fall across the other levels of schooling, but always remain significantly different than the excluded post-secondary class. In many cases we see no clear patterns that distinguish farmers in the middle three levels of schooling (literate to primary complete).

Table 29: Marginal Effects on the Probability of Choosing Each Category by Region

	Family		Credit		Technical Assistance		Specialization	
	Semi-Arid	South	Semi-Arid	South	Semi-Arid	South	Semi-Arid	South
Female	0.93	0.59	-0.51	-1.23	-0.79	-1.09	-0.46	-0.31
Level of schooling in relation to higher education								
Illiterate	1.52	2.10	2.26	0.80	-3.03	-3.53	1.18	1.88
Literate, but no schooling	1.37	1.98	2.39	0.64	-2.75	-2.77	0.95	1.06
Primary incomplete	1.21	1.88	2.58	0.77	-2.60	-2.57	1.03	1.21
Primary complete	1.17	1.63	2.19	0.76	-2.59	-2.48	0.53	0.73
Secondary complete	0.53	1.33	1.44	0.81	-2.72	-2.21	0.07*	0.22*
Years in charge of current farm in relation to more than 10 years								
0-1 years	-0.67	-0.80	-2.34	-1.66	-0.59	-1.78	-2.13	-2.17
1-5 years	-0.45	-0.44	-0.35	-0.53	0.03*	-0.54	-0.49	-0.49
5-10 years	-0.07*	0.15	-0.30	-0.21	0.15	-0.22	-0.34	-0.18
Age in relation to more than 65 years								
25-35	-1.48	-1.57	0.99	2.74	0.48	1.53	0.10*	-0.44
35-45	-1.40	-1.56	0.85	2.55	0.58	1.56	0.32	0.03*
45-55	-1.13	-1.46	0.33	2.36	0.46	1.20	0.54	0.24
55-65	-0.65	-0.80	0.48	2.16	0.40	0.86	0.56	0.50

Source: Agricultural Census-IBGE (2006)
Note: * = not significantly different from zero at 1%. All other effects are statistically significant at 1%.

5.5 Differences in Land Productivity: Credit, Technical Assistance and Specialization

Tables 30-32 show the decompositions in both regions for the cases of credit, technical assistance, and specialization. Table 30 shows log differences in land productivity between users and non-users of credit that are roughly twice as high in the South as in the Semi-Arid Northeast. In the Semi-Arid Northeast, in all cases other than the 100-500ha class, differences in the use of factors of production fully explain the differences in land productivity. There is nothing left over for TFP to explain. In the South, in contrast, yield differences are attributable to differences in both the use of inputs and TFP. The relative importance of the two varies with farm size. Interestingly, credit does not appear to be associated with the use of a superior technology. The TFP advantage of credit users is derived from non-factor variables.

Table 31 shows the decomposition for technical assistance. In the Semi-Arid, as with credit, the entire land productivity advantage of farmers with technical assistance is due to more intensive use of factors of production, not to TFP. The difference in TFP is negative in most cases—favoring those without credit—but small. In the South, in contrast, TFP differences account for around two-thirds of the yield advantage of farms with technical assistance, and differences in technology are the source of the TFP advantage. Two conclusions can be drawn. First, the importance of technology in explaining the TFP advantage among users of technical assistance in the South suggests that technical assistance is more associated with technical change than credit. Second, the fact that TFP differences play a role in explaining land productivity differences in the South, but not in the Semi-Arid, helps to explain why credit and technical assistance are associated with much lower levels of poverty in this region. Poverty was 50 percent lower in the South among users of credit or technical assistance, while it was only 10 and 20 percent lower, respectively, in the Semi-Arid Northeast.

Table 32 decomposes the log differences in land productivity between specialized and non-specialized farms. As we saw in previous tables, the differences in the use of factors of production are small. Here we see that the log differences also tend to be small, but they actually favor the non-specialized farms. As a result, the land productivity advantage of the specialized farms is attributable to differences in TFP. One way of interpreting this is that the specialized farmers produce higher value crops. The output advantage that is not attributable to a more intensive use of factors of production shows up as a TFP advantage.

Table 30: Land Productivity and Use of Credit

	Δ Type Conditional on Size				
	$y_c - y_{nc}$	dFat	dTec	dNfat	dTFP
	difference in logs				
Semi-Arid					
0-5	0.21	0.24	-0.29	0.25	0.04
5-20	0.21	0.23	-0.41	0.39	-0.02
20-100	0.13	0.13	-0.52	0.52	0.01
100-500	0.01	-0.16	-0.98	1.14	0.17
South					
0-5	0.45	0.18	-0.13	0.41	0.28
5-20	0.47	0.27	-0.32	0.52	0.20
20-100	0.47	0.32	-0.38	0.53	0.15

Source: Agricultural Census-IBGE (2006)

Table 31: Land Productivity and Use of Technical Assistance

	Δ Type Conditional on Size				
	$y_{ta} - y_{Nta}$	dFat	dTec	dNfat	dTFP
	difference in logs				
Semi-Arid					
0-5	0.39	0.43	-1.00	0.96	-0.04
5-20	0.34	0.45	-0.91	0.80	-0.11
20-100	0.28	0.39	-1.10	0.99	-0.11
100-500	0.29	0.28	-0.53	0.54	0.01
South					
0-5	0.65	0.18	0.39	0.08	0.47
5-20	0.66	0.26	0.32	0.08	0.39
20-100	0.65	0.27	0.31	0.08	0.39

Source: Agricultural Census-IBGE (2006)

Table 32: Land Productivity and Specialization

	Δ Type Conditional on Size				
	$y_{sp} - y_{Nsp}$	dFat	dTec	dNfat	dTFP
	difference in logs				
Semi-Arid					
0-5	0 1	-0.10	0.11	0.17	0.28
5-20	0.18	-0.11	0.10	0.19	0.29
20-100	0.15	-0.10	0.04	0.21	0.25
100-500	0.11	-0.11	0.03	0.19	0.22
South					
0-5	0.33	-0.06	.38	0.02	0.40
5-20	0.21	-0.15	0.44	-0.07	0.37

Source: Agricultural Census-IBGE (2006)

6. Conclusions

This Report explored the relationship between farm size, productivity and poverty in Brazilian agriculture. The Report began with a description of the characteristics of family farms, and how they differ from non-family farms. Evidence was provided on differences in poverty between family and non-family farms, and how they vary according to farm size, productivity, and off-farm income. Using a model that allows for differences in technology across farm size and type, econometric estimation and simulations were performed to shed light on differences in productivity and the mechanisms—technology, factor deepening, policies—from which they arise. The objective was to explore the extent to which enhanced agricultural productivity could bring about greater competitiveness for family farmers, increase their incomes and help to lift them out of poverty. We summarize key findings here, and then reflect on their importance for policy.

6.1 Findings

The first decade of the 21st century was good for Brazil. After a period of relative stagnation of real per capita income in the 1980s and 1990s, the country achieved real per capita growth above 3 percent per year since 2003. Simultaneously, improved social policies contributed to declining inequality since the late 1990s. Income growth and falling inequality combined to cut poverty by 40 percent and extreme poverty by around half in this period.

Positive trends in the Brazilian economy were spread broadly, but important differences persist between urban and rural areas, among macroregions of the country, and within agriculture. Although poverty has declined throughout Brazil, it remains much higher in rural areas, in the North and Northeast macroregions, and among people who work in agriculture. Extreme poverty is nearly five times as common in rural versus urban areas of Brazil, and exceeded 35 percent of the rural population in the North and Northeastern macroregions in 2010. Thus, of the more than 16 million people that the Federal government considered to be poor, close to half lived in rural areas.

The agricultural sector was booming in recent decades, with aggregate output growth rising by 3.77 percent annually since 1975, and with most growth coming from improvements in total factor productivity (TFP). Yet within the sector, an export-oriented commercial sector has benefited much more than most family farms. A small group of around 10 percent of the family farms increased its share of family farm output from around one-half in 1996 to two-thirds in 2006. This group is competitive and dynamic, and produces a decent standard of living for its family members. Over one-half of family farms, in contrast, generate little income. This Report sought to understand why.

One of the main conclusions of this Report is that family farms have a high rate of poverty in Brazil because many of them have insufficient land and because they produce with extremely low levels of productivity. In the Northeast, more than one-half of family farms have between zero and five hectares of land and 81 percent of them do not generate enough farm income to lift their full-time equivalent family labor above the poverty line. A much lower share of family farms are poor in the South. This reflects differences across regions in productivity and in the land distribution. In the South only 24 percent of family farms have zero to five hectares of land, and this group achieves higher levels of productivity when compared to the Northeast. Based on farm income alone, 62 percent of this group in the South is poor. For farms with a bit more land—between five and 20 hectares—poverty based solely on farm income falls to 35 percent in the South, but only to 67 percent in the Northeast. Poverty reduction among family farmers—especially in the Northeast—requires policies that address both insufficient land and low levels of productivity.

Within regions, the level of productivity is also an important determinant of poverty (based solely on farm income) for farms of each size, whether family or non-family. In the Northeast, and restricting attention to farms between five hectares and 20 hectares, 80 percent of farms in the bottom fifth of the land productivity distribution are poor, while 52 percent remain poor in the top fifth of the distribution (see Figure 7). In the South, poverty falls from 54 percent to 26 percent as we move from the bottom to the top quintiles of the land productivity distribution (see Figure 8). Again, productivity can make an important contribution to poverty reduction but—especially in the Northeast—poverty reduction requires multiple policies.

Low productivity is related to insufficient levels of physical capital, purchased inputs, and human capital. Huge differences were observed across regions in the use of capital on farms. In the South and Southeast, small farms (both family and non-family) with only zero to ten hectares of land had around R$50,000 of on-farm assets (see Table 12). In the Northeast, farms of this size in Alagoas and Pernambuco had only 10 to 15 percent of this level of capital. Within regions, there was also ample evidence of differences in the use of capital between family and non-family farms. For example, non-family farms in the South with five to 20 hectares of land used 66 percent more capital per hectare than did family farms in the same region. In the Semi-Arid Northeast, non-family farms used more than twice the capital per hectare used by family farms. Similar differences were observed in the use of purchased inputs within each region. As a group, non-family farmers were younger and more educated: 39 percent of family farmers in Brazil were at least 55 years old, and 85 percent had not even completed a primary education. Non-family farmers were 6 percentage points less likely to be over 55 and 23 percentage points more likely to have a complete primary education or higher. These differences in schooling and the use of capital and purchased inputs contributed to the divergence in productivity and income across farms.

Higher levels of schooling are correlated with increased use of credit and technical assistance, fertilizers, irrigation, and specialized production. The use of many of these items was also correlated with higher levels of productivity and lower levels of poverty. Farmers who used technical assistance achieved levels of land productivity that—depending on farm size—were about one-third higher in the Semi-Arid Northeast than those who did not, and about two-thirds higher in the South. Increased yields in the Semi-Arid were attributable to greater intensity in the use of factors, whereas increased yields in the South were related to both the use of factors and higher levels of TFP. While these relationships are not necessarily causal, farmers who used both credit and technical assistance, for example, had 20 percentage points less poverty in the Semi-Arid Northeast, and 39 percentage points less poverty in the South, than farmers who used neither.

We uncovered two striking stylized facts that merit further investigation. First, it is commonly accepted in Brazil that family farms have higher land productivity than non-family farms. Many argue that this relates to superior incentives that family members have for both work and supervision. While true in the aggregate, we argue that this is an artifact of differences in the size distribution of the two groups. Small farms have higher land productivity than large farms, and family farms tend to be small. When one compares farms of the same size, non-family farms uniformly have higher land productivity, although in the zero to five hectare class—where the distinction between family and non-family is less clear-cut—the difference might not be very large. The non-family farm advantage with regard to land productivity appears to be a result of a more intensive use of purchased inputs and capital, and higher levels of human capital, not of an inherent superiority of one type over another.

A second striking finding relates to the importance of off-farm work for non-family farmers. Non-family farms in the Northeast with zero to five hectares of land, for example, had the same rate of poverty as family farms (81 percent) when only agricultural income was considered. Yet off-farm income was extremely important for the non-family farmers. The inclusion of off-farm income lowered poverty by 28 percentage points, from 81 percent to 53 percent, for this group. For family farms, the poverty rate only dropped by 2 percentage points when off-farm income was included. The same phenomenon was observed in the South, where the inclusion of off-farm income lowered poverty by 32 percentage points for non-family farms, but by only 5 percentage points for family farms. Thus, non-family farmers not only hire labor in (one of the criteria for distinguishing between family and non-family farms), but some of them also hire out a significant amount of their own labor. The non-family farmers appear to have more education and skills, and thus are more competitive in labor markets. We suspect that they use this advantage to generate cash, relax credit constraints, and permit operating their farm at a higher level of productivity than family farmers. Thus, off-farm employment might contribute to poverty reduction both through increasing income directly, and by permitting higher levels of capital and purchased inputs to boost agricultural income on their own farms.

The result above speaks more to the importance of off-farm income than to the distinction between family and non-family farms. The family/non-family dichotomy is less clear-cut for small farms, and needs to be qualified by certain limitations of the legal definition of a family farm. A weakness of the legal classification, especially at the lower end of the farm size distribution, is that farms may be classified as non-family solely because they are well-inserted into the labor market. In our view, farms do not cease to be family farms when they earn 51 percent of their income off-farm. While this is a limitation of the legal definition, it does not undermine the findings about the importance of off-farm income for poverty reduction among small farmers.

While many family farms have insufficient land, and lower levels of land productivity conditional on farm size, we conclude that they are "efficient but poor" (Schultz, 1964). On average, they appear to be using efficiently the little that they have. In both regions of the country we observed that non-family farms use inputs more intensively, and have advantages with non-factor inputs such as human capital or climate and soil, but family farms off-set these disadvantages with higher levels of TFP. Thus, they close some of the yield gap by doing a better job at turning inputs into outputs. Nevertheless, with limited land and low levels of capital and purchased inputs, many are unable to generate sufficient income to escape poverty.

6.2 Policy Implications

The "efficient but poor" hypothesis raises the question of the best way to increase income for poor family farmers. Consistent with World Bank (2003), *World Development Report 2008* and Helfand and Pereira (2012), we believe that there are numerous pathways out of rural poverty, and that multiple policies are necessary to assist the poor to make successful transitions from poverty. The pathways include: (i) agriculture—either through intensification of family farms or wage labor; (ii) non-agriculture—either through labor market earnings or self-employment; (iii) migration, for those households that choose to exit from the sector; and (iv) transfers for those households without the potential to generate sufficient earned income.

This Report focused on intensification of family farms, without ignoring the importance of the other exit pathways, and with full awareness that the agricultural pathway is only appropriate for a minority of family farms. Where producers have sufficient land, poverty reduction depends on increasing productivity and income. In both regions that were studied, farms that used credit, technical assistance, irrigation or that specialized in production, often generated two to three times the profit per family member of farms of the same size that did not do so. As a result, poverty was significantly lower for these farms. In both regions, technical assistance was more strongly associated with gains in land productivity and income than was credit. This suggests that while credit can relax constraints and permit increased use of purchased inputs, technical assistance is more closely related to technical change and income growth. This finding reinforces the importance of investments in technical assistance to spur needed technology adoption among family farms as a means of stimulating TFP growth. Similarly, incentives to adopt new technologies can be increased by improved insurance mechanisms. This is especially important in risky climactic environments such as the Semi-Arid portion of the Northeast.

For farms with insufficient land, family farming can only provide a pathway out of poverty if access to additional land is provided. Thus, land reform can be one piece of the solution, especially in the Northeast where over one-half of the farms have less than five hectares. The results of this Report suggest, however, that family farmers require much more than land in order to escape poverty. Land reform programs must either be located in regions where off-farm labor opportunities can complement farm income, or they must target an adequate farm size combined with high enough levels of productivity to avoid the reproduction of poverty. In addition to redistributive land reform, a policy of defragmentation that could permit creating family farms of an appropriate size where *minifundia* currently exist should be a high priority. In this regard, expanding the scope of the market-assisted land reform program *Crédito Fundiário* would hold considerable promise. A more active land market, especially in the Northeast, requires clarifying property rights and providing land titles where they are lacking.

Policies that support access to off-farm earnings can play an important role in poverty reduction. In both regions that were studied, off-farm earnings reduced poverty for small non-family farmers by more than off-farm earnings and transfers combined for family farms of the same size. Numerous studies have shown that access to off-farm employment increases with human capital. With regard to non-agricultural employment, proximity to population centers is another important factor that increases the probability of finding employment. There are many possible growth engines that can create labor demand for the poor. In some regions, the existence of irrigated agriculture might create linkages to non-agricultural jobs in fruit and vegetable processing. In other locations it might be an abundance of animal production that creates employment in a slaughter house. The diversity of possibilities for growth and employment has led naturally to a focus on territorial development. If alternatives to migration are to be constructed, local territories must find their own dynamic sectors that have the potential to generate employment. It is not essential that these sectors be based on agriculture. What matters is that they create opportunities that are accessible to the poor.

Among the highest priorities for public policy in Brazil should be to improve the quantity and quality of education for young people who live in rural areas. This is perhaps the only policy that contributes positively to all pathways out of poverty. Education is associated with higher agricultural income as a result of its relationship with productive efficiency, technological adoption, and the ability to participate in input and output markets. Education is also associated with higher non-agricultural income because it increases the probability of finding employment and the earnings of individuals once employed. Education is an important factor that contributes to more successful migration histories. Finally, education is a key component throughout the world in cash transfer programs that seek to break the transmission of poverty from one generation to another. In spite of these benefits, education for rural households continues to lag behind urban areas in both quantity and quality. The international literature has shown convincingly that the social returns to investments in education—especially at the primary level—are quite high. In order to reap the full long-term benefits of *Bolsa Família*, and provide a more promising future for rural youth regardless of the pathway that they pursue, Brazil needs to improve the quality of schools for its rural population.

It is often observed that farmers are not young in Brazil. While it is true that 39 percent of family farmers were over 55 years old in 2006, it is also true that 34 percent were between 25 and 45. The younger cohorts have more schooling, and appear to be more willing to experiment with new technologies. In order for agriculture to provide a pathway out of poverty for this group, it is essential that policies assist them to produce with sufficient land and at much higher levels of productivity than the previous generation of farmers. Nearly 50 years after Schultz's seminal publication, it is time to finally achieve "efficient and not poor."

7. Technical Annex

7.1 The heterogeneous translog production function

The standard stochastic frontier production function assumes a common technology – available to all producers – that describes the production possibilities frontier.[14] There are random and unobserved factors that make a stochastic model, with the inclusion of a random error, preferable to a deterministic model of production such as Data Envelopment Analysis (DEA). In addition, there are many reasons why a producer might not operate on the frontier of what is technologically possible, and these factors are captured by the one-sided error term that is added to the production function to capture such idiosyncratic behavior. The one-sided error measures the technical inefficiency of each producer.

Thus, in the case of a translog functional form, the standard model is

$$y = \alpha + x\beta + x'x\gamma + z\delta - u + v \quad u\sim N^{+}(0, \sigma_u^2) \;\; v\sim N(0,\sigma_v^2) \tag{1}$$

where *y* is the log of a scalar output, *x* is a vector of inputs in logs, *x'x* are the cross and quadratic terms of the translog, *z* is a vector of non-factor variables that condition the production function, *u* is the one-sided inefficiency term, *v* is a random error assumed to be normally distributed with mean zero and variance σ_v^2, and α, β, γ and δ are technology parameters that need to be estimated. There are many possible asymmetric distributions that could be appropriate for *u*. The most common in the literature is a half-normal distribution truncated at zero with variance σ_u^2.

If the asymmetric inefficiency term (*u*) is small relative to the random error (*v*), it might not be possible to separate the two components of the error. In this case, the model collapses to an average—rather than a frontier—translog production function. This is the case here. We believe that this happens in our study for this Report because we do not have farm-level data. As mentioned in the introduction, and explained in more detail in Section 4, we use aggregate data on "representative farms". These reflect the average in each municipality for each class (i.e., family, non-family) and size (i.e., four categories in terms of hectares). Because the data are

[14] Useful reviews of the stochastic frontier literature can be found in Coelli et al. (1998), Kumbhakar and Lovell (2000) and Greene (2007).

aggregated, it is likely that positive and negative values of the idiosyncratic component for each individual farmer—which captures farm-level inefficiency—cancel each other out. Thus, the one-sided inefficiency term *u* drops out of the model.

The standard model above assumes that there is a common technology. However, farms are (i) located in regions that have different conditions and potential; (ii) are managed by people with differing degrees of technical knowledge and experience; and (iii) are inserted in economic environments with many different constraints. These factors create the possibility that producers might choose different technologies. Much of the stochastic frontier production function literature does not explicitly address this type of technological heterogeneity, thus including it in the one-sided error term that captures each producer's "inefficiency."

There is, however, a strand of the literature that is concerned with this sort of heterogeneity. Studies of the adoption rates of new technologies highlight the importance of differences in the technologies that are actually utilized. Mundlak (1988, 1993, 2001) has written extensively about the endogenous "implemented technologies" that producers choose, suggesting that there are observable "state variables" that influence these choices. Larson and León (2006) apply such a model in the context of Ecuadorean agriculture. They allow for a continuum of technologies that depend on a set of state variables. The state variables influence the level of technology (α) and the elasticities of output with respect to inputs. State variables could also be posited as influencing the variance of the inefficiency term.[15]

Orea and Kumbhakar (2004) and Greene (2003, 2005) pursue a different approach to technological heterogeneity. They develop latent class models that allow for a set of alternative technologies, but assume that we can only know the probability with which a given producer uses each technology. Thus, the problem becomes one of jointly estimating the probabilities with which each technology is used and the parameters of each production function. Alvarez (2010) provides an interesting application for Spanish dairy farms, and Wollni and Brümmer (2012) use a similar approach to analyze coffee producers in Costa Rica. We follow in the spirit

[15] The state variables used by Larson and León fall into four broad categories: farmer characteristics (formal education, agricultural education and gender), social capital (affiliation with producer association, indigenous language use), markets and institutions (private markets for credit and technical assistance, intermediate buying arrangements, and participation in output markets, distance to markets), and nature and risk (climate and topology measures).

of these approaches and use the Heckman (1979) sample selection model to correct for the selection bias that would result from estimating the production function unconditionally.

The approach assumes that the producers in each class (*c*) are heterogeneous, and that belonging to a particular class is an endogenous decision that depends on a vector of variables (*w*):

$$\text{Prob (c)} = f\,(\text{w}) \tag{2}$$

The probability of belonging to the class c=1 (family) rather than c=0 (non-family) can be estimated with a probit. The estimates are used to construct the inverse Mills ratio (λ) which is included as an additional variable in the production function. Failure to control for the selection mechanism, and the possibility that unobserved variables influence both selection and production, could lead to inconsistent estimates of the coefficients in the production function.

In addition to the inclusion of λ, equation (1) is modified to incorporate heterogeneity in the technology across classes (*c*) and sizes (*k*). In order to limit the number of coefficients that need to be estimated, and reduce problems of multicollinearity among the regressors, we assume complete heterogeneity between the classes *c* (i.e., family, non-family), but restrict the degree of heterogeneity across sizes (*k*). We assume that the heterogeneity across farm sizes can be captured by the intercepts and linear terms of the production function. This implies that: (i) the quadratic and cross terms in the translog; and (ii) the influence of the non-factor variables (*z*) are identical for all producers in each class (*c*). Based on these assumptions the model becomes:

$$y = \alpha(t) + x\beta(t) + x'x\gamma(c) + z\delta(c) + v \qquad v \sim N(0, \sigma_v^2(c)) \tag{3}$$

where t=(c,k).

A final assumption which draws on a large body of evidence in development and agricultural economics (see Mundlak 2001) is that returns to scale are approximately constant in agriculture. Based on this literature, we assume a weak version of constant returns to scale (CRS). We impose CRS within each of the four size classes, but allow the technology to have variable returns across them. The imposition of CRS implies that output (*y*) and the inputs (*x*) are all divided by area. Thus, what matters are the input and output intensities (such as, for example, labor per hectare).

The imposition of CRS has an additional benefit. It provides us with a variable—farm size—that is a clear candidate for identification of the inverse Mills ratio. Identification requires finding at least one exogenous variable that influences the probability of selection, and thus belongs in the probit, but does not enter the production function given in equation (3). Similar to the state variables discussed earlier, we include a vector of exogenous variables (*w*) in the probit. The w vector includes (i) temperature; (ii) rainfall; (iii) distance to markets; and (iv) age and education of the producer, which are likely to influence the selection decision. However, many of these variables also enter the production function via the vector (*z*). In contrast, the scale of production as measured by farm size has no place in the production function. It should certainly influence the probability of hiring labor, and thus of becoming a non-family farmer by definition, but with the assumption of CRS there is no reason for farm size to be included in the production function.

7.2 Counterfactual Simulations of Productivity across Farm Types and Sizes

The heterogeneous production function specified in equation (3) will be used to conduct counterfactual simulations to address what would happen to the productivity of family farms under a variety of scenarios. We conduct two types of experiments. The first uses a Oaxaca-Blinder approach to decompose the differences in productivity between the mean of each type (*t*) into components attributable to differences in: (i) technology; (ii) non-factor variables; and (iii) factors of production. The sum of the first two components can be interpreted as the difference in total factor productivity (TFP) between any two groups. The second exercise is conducted for all "representative farms" in the data. For each of these farms, we simulate the productivity gain of using a different technology. The simulations can be conducted unconditionally or conditional on farm size or farm class. The results include average productivity gains for each group, as well as the distribution of the choices that optimize the gains.

The first exercise focuses on the mean of each type (*t*), where *t* has eight elements that vary by class (i.e., family, non-family) and by farm size (i.e., 0-5ha, 5-20ha, 20-100ha, and 100-500ha). In order to illustrate the approach, we rewrite equation (3) as:

$$y_t = \alpha_t + \phi_t f_t + \delta_t z_t \qquad (4)$$

where y is output per unit of land, f represents the linear and cross terms of the inputs in the production function (x, xx), z is a vector of non-factor variables that influences production, and α, $\phi=(\beta,\gamma)$ and δ are the estimated intercepts, technology coefficients, and coefficients on non-factor variables.[16] The t subscripts capture the heterogeneity of the production function across the eight types of producers.

The Oaxaca-Blinder decomposition between types t=a,b can be written as follows:

$$y_a-y_b = \{\alpha_a-\alpha_b+(\phi_a-\phi_b)(f_a+f_b)/2+(\delta_a-\delta_b)(z_a+z_b)/2\} +\{(z_a - z_b)(\delta_a+\delta_b)/2\} + \{(f_a- f_b)(\phi_a+\phi_b)/2\}$$

which is the sum of three components:

1) the difference in technologies (dtec) $=\{\alpha_a-\alpha_b+(\phi_a-\phi_b)(f_a+f_b)/2+(\delta_a-\delta_b)(z_a+z_b)/2\}$
2) the difference in non-factor variables (dnfat) $= (z_a-z_b)(\delta_a+\delta_b)/2$
3) the difference in factors of production (dfat) $= (f_a-f_b)(\phi_a+\phi_b)/2$

The difference in total factor productivity is dtfp = dtec + dnfat. Rather than choose one of the two groups as a reference for the simulation, we evaluate the difference in technologies at the average of the factor $[(f_a+f_b)/2]$ and non-factor variables $[(z_a+z_b)/2]$, and we evaluate the difference in factors of production and non-factor variables at the average of the technology coefficients $[(\phi_a+\phi_b)/2]$ and $[(\delta_a+\delta_b)/2]$.

By focusing on the mean of each group, the Oaxaca-Blinder approach ignores the effect of heterogeneity with each group. For this reason, we conduct a complementary exercise that examines the importance of differences in the factors of production and non-factor variables across producers within each of the eight types. For each representative farm we examine the effect on productivity of a change in technology. We do this by identifying the technology that maximizes productivity, given the unique vector of inputs and non-factor variables that each representative farm utilizes. The simulation proceeds as follows. Let y(x|t) be the production frontier of the farm that utilizes the technology associated with each value of t. We consider three optimization scenarios.

[16] Because of the imposition of CRS, y and x have been normalized by the quantity of land (area). Thus, the dependent variable can be interpreted as land productivity (y/area) and the inputs can be interpreted as intensities (x/area).

- In the first, we assume that each representative farm chooses from all eight technologies (t) that are available in order to maximize productivity: $t^* = \text{argmax}\ (y(x|t);\ t \in T)$.
- The second scenario restricts the choice set T to those technologies that are used by the same class (c) of farms, such as family farms: $t^*|c = \text{argmax}\ (y(x|t);\ t \in T(c))$. For each class ($c$), there are four feasible technologies available.
- The final scenario restricts the set of technologies T to those used by farms of the same size (k): $t^*|k = \text{argmax}\ (y(x|t);\ t \in T(k))$. For each farm size there are only two feasible technologies in the restricted set T(k).

With this approach, we can calculate three measures of inefficiency for each representative farm:

- $U_i = y(t^*) - y_i$
- $U_i|c(i) = y(t^*|c_i) - y_i$
- $U_i|k(i) = y(t^*|k_i) - y_i$

These measures can be used to calculate the distribution of optimal choices under each scenario, or the average gain for each type (t) from using the technology that maximizes productivity. Note that the average gain is calculated from the distribution of individual gains, not from the mean gain as with the Oaxaca-Blinder decomposition.

7.3 The Data: Representative Farms as a Unit of Analysis

It is important to emphasize that the models discussed above were not estimated with farm-level data. Due to issues associated with the confidentiality of the microdata, all of the data used were aggregated by municipality, farm size (0-5ha, 5-20ha, 20-100ha, and 100-500ha), and class (family, non-family). Data aggregation implies that we assume homogeneity within each aggregate observation, for example family farms with five to ten hectares in the municipality of Viçosa, MG. We call these "representative farms." The econometric analysis will explore variance between representative farms, but due to the aggregation, cannot examine variance within them (see Box 3 earlier in this Report for a detailed discussion of representative farms). This is an additional reason for emphasizing technological heterogeneity rather than idiosyncratic inefficiency, as the latter is specific to the individual farm. In any case, the data averaging resulting from aggregation reduces the importance of idiosyncratic inefficiency.

Prior to aggregating the data into representative farms by municipality, size, and class, we filtered the data to remove spurious observations. When working with a preliminary dataset, we observed unbelievably large values of land productivity and other variables. For this reason, and after some sensitivity analysis, we elected to filter the data as follows. We defined five variables that were important for our study, and removed the largest one-half of one percent (0.5%) of the observations for each variable in each state. The variables, all defined per hectare, were: (i) value of output; (ii) value of expenditures on inputs; (iii) value of the capital stock; (iv) value of sales; and (v) total number of family members working on the farm. Thus, by way of example, in the state of Pernambuco, where there were 304,788 farms, we removed 5,094 observations (or 1.67 percent of the total) prior to aggregation. This was somewhat less than 2.5 percent (5 exclusions x .05 percent each) because the filters were applied simultaneously and some observations fell into more than one filter.

One of the limitations of working with aggregate data is that, in order to protect the confidentiality of the data, IBGE censors the value of a variable that has fewer than three observations reporting information. Thus, if there were five family farms between 100 and 500 hectares in the municipality of Água Branca, Alagoas, but only two of them report having family members who worked on the farm less than 60 days, the value of this variable would appear as missing. In order to reduce the loss of representative farms, we decided to impute data when we had additional information and in cases that we were reasonably comfortable that the imputation would not affect the results. Thus, because in the example above we know that each farm had to have had at least one person working on the farm less than 60 days, we imputed the lower bound of one person per farm. In other cases, for example, if "unusable land" was missing and we needed to subtract this component from total area in order to calculate net area, we imputed the average unusable share of total area for that farm type, size, and region. In no case did we impute values for key variables such as total value of output or total amount of land. Because cells with few farms are the ones that get censored, this implies that we lose a higher share of representative farms (cells) than actual farms.

The econometric analysis was conducted for the South macroregion and the Semi-Arid portion of the Northeast macroregion. Table 33 shows the number of representative farms, and individual farms, of each size and class in the Semi-Arid and South. We estimate the models in each region with around 7,000 observations (7,144 in the Semi-Arid and 6,821 in the South) which aggregate the information from 1.481 million farms in the Semi-Arid and 926,000 farms in the South, which together account for nearly 50 percent of all farm establishments nationwide. On average, there are around 45 farms in each representative non-family farm in the Semi-Arid, and around 38 in the South. Representative family farms have around 350 and 240 farms contained in each observation in the Semi-Arid and South, respectively, although there are far fewer in the 100ha-500ha class.

There are at least 671 representative farms in every size and class, with the exception of family farms between 100ha and 500ha in the South. There are only 51 observations of this type, which generated imprecise and unstable econometric estimates. For this reason we exclude these large

Table 33: Representative Farms

	Semi-Arid			South		
	Representative farms	Farms (1000s)	Farms / Representative farms	Representative farms	Farms (1000s)	Farms / Representative farms
Total	7,144	1,481	207	6,821	926	136
Family (F)	3,796	1,330	350	3,282	790	241
0-5	964	705	732	1,054	190	180
5-20	988	346	350	1,073	380	354
20-100	1,027	247	240	1,104	218	198
100-500	817	32	39	51	1.9	36
Non-Family (NF)	3,348	151	45	3,539	136	38
0-5	671	54	81	675	19	28
5-20	814	32	39	924	29	31
20-100	922	34	37	1,023	39	38
100-500	941	31	33	917	49	54

Source: Agricultural Census-IBGE (2006)

family farms in the South from the simulations. In the Semi-Arid region, we lost 20 percent of the representative farms due to aggregation and confidentiality, but only 7 percent of the individual farms. In the South we lost 21 percent of the representative farms, but only 3.4 percent of all farms.

Box 4 earlier in this Report defines the variables that compose the heterogeneous translog production function used in the model estimation. We consider the variables in *g* and *p* to be either exogenous or determined before the current agricultural cycle. Among the inputs (*x*), the endogeneity of purchased inputs is what most concerns us. Nor can one consider the use of credit, technical assistance, electricity in production, or irrigation as exogenous. These are clearly endogenous variables as the 2006 Agricultural Census does not measure the *availability* of these factors but rather their *use*, reflecting the producer's choice to borrow or not, contract technical assistance or not, etc. Thus, we instrument for purchased inputs and the policy-related variables (*w*) with a fixed effect for each microregion, farm size dummies, all of the variables in *g* and *p*, and variables that capture the number of private bank branches and public bank branches in each municipality in the period 2000-2005. In all cases, F-tests strongly rejected the presence of weak instruments.[17]

[17] The values of the F-statistics range from 13 in the case of w_2 in the Semi-Arid to 204 for x_2 in the Semi-Arid. The tests are conducted with 47 parameters, 7,144 observations in the Semi-Arid, and 6,821 in the South.

8. Bibliography

Alston, Julian M., Babcock, Bruce A., and Pardey, Philip G. (Eds.). *Shifting Patterns of Agricultural Production and Productivity Worldwide.* The Midwest Agribusiness Trade Research and Information Center Iowa State University, Ames, Iowa, 2010.

Alvarez, Antonio and Julio del Corral, "Identifying different technologies using a latent class model: extensive versus intensive dairy farms," *European Review of Agricultural Economics,* 37:2, 2010 pp. 231–250.

Banco Central do Brasil. *Anuário Estadístico de Crédito Rural.* 2011

Butzer, R., Y. Mundlak, and D. Larson, "Measures of Fixed Capital in Agriculture," Policy Research Working Paper 5472, World Bank, November 2010.

Coelli, T. P. Rao and G. Battese, *An Introduction to Efficiency and Productivity Analysis,* Boston: Kluwer Academic Publishers, 1998.

Coelli, Tim and Euan Fleming, "Diversification economies and specialisation efficiencies in a mixed food and coffee smallholder farming system in Papua New Guinea," *Agricultural Economics* 31, pp. 229–239, 2004.

De Negri, João Alberto and Kubota, Luis Claudio (Eds.). *Políticas de Incentivo à Inovação Tecnológica*. Brasília. 2008.

Filho, José Eustáquio Ribeiro, "Radiografia Produtiva e Tecnológica da Agricultura Familiar no Brasil," Mimeo, May 28, 2012.

Food and Agricultural Organization. *How to Feed the World in 2050*. FAO:Rome. 2011

França, Caio Galvão de, Mauro Eduardo del Grossi, Vicente P.M. de Azevedo Marques, *O Censo Agropecúario 2006 e a Agricultura Familiar no Brasil,* Brasília: MDA, 2009.

Gasques, José G., Eliana T. Bastos, Constanza Valdes, and Mirian R. Bacchi. Produtividade da Agricultura Brasileira e os Efeitos de Algumas Políticas**.** 2012

Global Harvest Initiative. *Global Gap Report: Measuring Agricultural Productivity.* 2011

Greene, William H., "Estimation of a Latent Class Stochastic Frontier Model," manuscript, Department of Economics, Stern School of Business, New York University, 2003.

Greene, William H., "Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model," *Journal of Econometrics*, 126, pp. 269-303, 2005.

Greene, William H., “The Econometric Approach to Efficiency Analysis” in Fried, H., K. Lovell and S. Schmidt, *The Measurement of Productive Efficiency and Productivity*, Oxford University Press, 2007.

Guanziroli, Carlos E., Alberto Di Sabbato, Antônio M. Buainain, “Dez Anos de Evolução da Agricultura Familiar no Brasil: 1996 e 2006,” mimeo, 2010.

Guanziroli, C. E. “Pronaf Dez Anos Depois: Resultados e Perspectivas para o Desenvolvimento Rural.” *Revista de Economia e Sociologia Rural*, v. 45, p. 301-328, 2007.

Guanziroli, C. A. Romeiro, A.M. Buainain, A. Di Sabbato, and G. Bittencourt, *Agricultura Familiar e Reforma Agrária no Século XXI*, Rio de Janeiro: Editora Garamond Ltda., 2001.

Helfand, Steven M., Ajax R. B. Moreira**,** Adriano M. R. Figueiredo, “Explicando as Diferenças de Pobreza entre Produtores Agrícolas no Brasil: Simulações Contrafactuais com o Censo Agropecuário 1995-96” *Revista de Economia e Sociologia Rural*, Vol. 49, No. 2, pp. 391-418, 2011.

Helfand, Steven, and Vanessa da Fonseca Pereira, “Determinants of Rural Poverty and Implications for Public Policy in Brazil” in Portuguese in *A Nova Cara da Pobreza Rural no Brasil: Transformações, Perfil e Desafios para as Políticas Públicas*. Editors: Buainain, Antônio Marcio, Cláudio Dedecca, and Henrique Neder. Brasilia: Inter-American Institute for Cooperation on Agriculture (IICA), 2012.

Helfand, S., Rocha, R., and VINHAIS, H. E. F., “Pobreza e Desigualdade de Renda no Brasil Rural: Uma Análise da Queda Recente,” *Pesquisa e Planejamento Econômico*, Vol. 39, No. 1, 2009, pp. 67-88.

Instituto Brasileiro de Geografia e Estatística, (IBGE), *Censo Agropecuário 1995-96*. Rio de Janeiro: IBGE, 1998.

Instituto Brasileiro de Geografia e Estatística, (IBGE), *Censo Agropecuário 2006: Brasil, Grandes Regiões, e Unidades da Federação* . Rio de Janeiro: IBGE, 2009.

Instituto Brasileiro de Geografia e Estatística, (IBGE), *Censo Agropecuário 2006: Agricultura Familiar, Primeiros Resultados*. Rio de Janeiro: IBGE, 2009.

Instituto Brasileiro de Geografia e Estatística, (IBGE), *Censo Demográfico 2010*. Rio de Janeiro: IBGE, 2011.

Jonasson, E., Helfand, S. “How Important are Locational Characteristics for Rural Non-agricultural Employment: Lessons From Brazil,” *World Development*, Vol. 38, No. 5, pp.727-741, 2009.

Kumbhakar, S. and K. Lovell, *Stochastic Frontier Analysis*, Cambridge: Cambridge University Press, 2000.

Kumbhakar, Subal C, Efthymios G Tsionas and Timo Sipilainen, “Joint Estimation of Technology Choice and Technical Efficiency: An Application to Organic and Conventional Dairy Farming,” *Journal of Productivity Analysis*, 31:3, pp. 151-61, June 2009.

Larson, Donald F. and Mauricio León, “How Endowments, Accumulations, and Choice Determine the Geography of Agricultural Productivity in Ecuador,” *World Bank Economic Review,* 20:3, pp. 449-71, 2006.

Mattei, L. *PRONAF 10 Anos: Mapa da Produção Acadêmica*. NEAD Estudos 12. Brasiília: MDA, 2009.

Mia, Irene et al. *The Brazil Competitiveness Report 2009.* World Economic Forum: Geneva. 2009.

Moreira Ajax R. B., Steven M. Helfand, Adriano M. R. Figueiredo, “Explicando as Diferenças na Produtividade Agrícola no Brasil,” Texto para Discussão No. 1254, Rio de janeiro: IPEA, Jan. 2007.

Mundlak, Yair, ‘‘Endogenous Technology and the Measurement of Productivity,’’ in S. M. Capalbo and J. M. Antle, eds., *Agricultural Productivity: Measurement and Explanation*, Washington: Resources for the Future, 1988.

Mundlak, Yair ‘‘On the Empirical Aspects of Economic Growth Theory,’’ *American Economic Review Papers and Proceedings*, 83:2, pp.415–20, May 1993.

Mundlak, Yair, ‘‘Production and Supply,’’ in B. Gardner and G. Rausser, eds., *Handbook of Agricultural Economics. Volume 1A: Agricultural Production*, Amsterdam, London and New York: Elsevier Science, 2001.

Navarro, Zander. “A Agricultura Familiar no Brasil: Entre a Política e as Transformações da Vida Econômica,” chapter 7 in *A Agricultura Brasileira: Desempenho, Desafios, e Perspectivas,* edited by Gasques, J.C., J.E.R.V. Filho, and Z. Navarro, *Brasília:* IPEA, 2010, pp. 185-209.

OECD. *Fostering Productivity and Competitiveness in Agriculture*. OECD Publishing. 2011.

Orea, Luis and Subal C. Kumbhakar, “Efficiency Measurement Using a Latent Class Stochastic Frontier Model,” *Empirical Economics*, 29, pp. 169-83, 2004.

Osorio, Rafel Guerreiro, Pedro H.G.F. de Souza, Sergei S.D. Soares, and Luis F.B. de Oliveira, *Perfil da Pobreza no Brasil e Sua Evolução no Período 2004-2009.* Texto para Discussão 1647, Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília, August 2011.

Rios, Ana R., William A. Masters and Gerald E. Shively, “Farm Productivity and Household Market Participation: Evidence from LSMS Data,” paper prepared for the 27th International Conference of Agricultural Economists, Beijing, China, 16-22 August, 2009.

Schultz, T.W. *Transforming Traditional Agriculture,* New Haven: Yale University Press, 1964.

World Bank. *Country Partnership Strategy for the Federative Republic of Brazil 2012-2015.* Washington, DC: World Bank. 2011

World Bank, The World Development Report 2008: Agriculture for Development, Washington D.C.: World Bank, 2007.

World Bank, Rural Poverty Alleviation in Brazil: Toward an Integrated Strategy, Washington D.C.: World Bank, 2003.