The Inverse Relationship between Farm Size and Productivity: Refocusing the Debate[†]

Steven M. Helfand (Corresponding Author)
Department of Economics, University of California, Riverside, CA 92521
steven.helfand@ucr.edu

Matthew P.H. Taylor
Department of Economics, University of California, Riverside
mptaylor1@live.com

This draft completed on June 24, 2019
Second draft completed on July 10, 2018
First draft presented at WEAI, July 2016

Abstract

The relationship between farm size and productivity is a recurrent topic in development economics, almost as old as the discipline itself. This paper emphasizes the importance of choice of productivity measures in the inverse relationship literature. First, we seek to clarify the common measures, their relationships, and advantages and limitations in empirical work. Second, we argue that total factor productivity (TFP), not land productivity, is the appropriate indicator for most policy questions. Third, we discuss the changing nature of the farm size – productivity relationship and show that the identification of these dynamics depends on the choice of measure. Lastly, using a pseudo-panel of Brazilian farms that are aggregated at the municipality and farm size levels over the period 1985-2006, we provide new evidence on the inverse relationship between farm size and both land productivity and TFP. The inverse relationship between size and land productivity is alive and well. The relationship between TFP and size, in contrast, has evolved with modernization during this period, becoming increasingly U-shaped or even positive. Policy implications are discussed.

Keywords: Inverse relationship, agriculture, farm size, total factor productivity (TFP), Brazil.

JEL Codes: O13, Q12

I. Introduction

The relationship between farm size and productivity is a recurrent topic in development economics, almost as old as the discipline itself. John Stuart Mill observed an inverse relationship as early as 1848, later positing that this had changed due to increasing capital intensity of farming (Lipton, 2009). The issue appeared in the works of Marx, resurfaced with Lenin and Chayanov in the early 20th century, and has captivated modern agricultural and development economists for over fifty years. Debate around the nature and causes of this relationship continues despite a mountain of empirical analysis, posing a puzzling question for 21st century researchers (Binswanger et al., 1995; Eastwood et al., 2010). Conventional economic wisdom expects resources to be allocated such that returns to land are equalized across farms; however, the empirical research on developing countries contradicts this and frequently identifies an inverse relationship. Policy-makers in developing countries have engaged the debate, as an inverse relationship between farm size and productivity indicates a role for small farms in development strategies and the potential for land reform to simultaneously generate improvements in equity and efficiency.

Harnessing such a relationship to inform policy requires correct interpretation of the empirical evidence as well as an understanding of its causes, the channels through which it operates, and the factors that condition its strength. Theoretical explanations for this phenomenon often result from household heterogeneity and/or (multiple) market failures, for example Sen's (1966) dual labor market hypothesis, Eswaran and Kotwal's (1986) model of household endowments with credit constraints, and Feder's (1985) model of moral hazard and costly monitoring of hired labor. Risk aversion (Barrett, 1996) and agronomic issues (Bevis and

1

Barrett, 2017) provide alternative explanations. Measurement error (Lamb, 2003; Carletto et al., 2013; Gourlay et al., 2017; Desiere and Jolliffe, 2018; Abay et al., 2018) and omitted variables, such as soil quality (Bhalla and Roy, 1988; Benjamin 1995; Assunção and Braido, 2007), are two empirical issues that could lead to an observed inverse relationship. Attempts to sort out the relative importance of these mechanisms have been mixed.

Adding to the confusion is the variety of productivity measures and empirical approaches that have been used. As with Sen (1962), Deolalikar (1981), Assunção and Braido (2007), Henderson (2015), Deininger et al. (2018), and Dillon et al. (2016), much of the early literature used land productivity—output per unit of land—as a measure of performance.[1] Conditioning land productivity on input use by estimating a production function is a second commonly used approach that generates an alternative measure of performance (Bardhan, 1973; Carter, 1984; Barrett et al., 2010; Ali and Deininger, 2015; Muyanga and Jayne, 2016). Controlling for a partial set of inputs (Bhalla and Roy, 1988; Desiere and Jolliffe, 2018) is distinct from estimating a full production function. Still others employ value added per unit of land (Heltberg, 1998; Carletto et al., 2013), profit per unit of land (Heltberg, 1998; Foster and Rosenzweig, 2017), profit (Benjamin, 1995; Lamb, 2003; Ali and Deininger, 2015), or technical efficiency (Helfand and Levine, 2004; Kagin et al., 2016). Despite the recognition that partial measures such as land productivity are problematic (Berry and Cline, 1979; Binswanger et al., 1995; Muyanga and Jayne, 2016), they continue to be used, often alongside alternative productivity measures, and are frequently discussed synonymously with a more general notion of productivity. Where multiple productivity

---

[1] The literature often uses yield and land productivity interchangeably. We only use yield when talking about a physical measure of productivity for a single product (tons/hectare). Land productivity is more appropriate in a multiple-output context, requiring a method for aggregation.

measures are used, the distinction between the relationships being estimated are seldom addressed directly. Conceptual clarity is needed on how these measures relate to each other and to farm size.

We show that an inverse relationship between farm size and a partial productivity measure, such as land productivity, is neither necessary nor sufficient for an inverse relationship between farm size and a comprehensive measure of productivity, such as total factor productivity. As such, these measures are not generally comparable. An inverse relationship may be observed when using land productivity, but not necessarily when using a comprehensive and preferable measure of productivity. In fact, Bardhan (1973), Berry and Cline (1979), Carter (1984), and Heltberg (1998) are all examples where an alternative productivity measure leads to an attenuated, if not direct, relationship between farm size and productivity, even though an inverse relationship holds using land productivity. This highlights the importance of how productivity is measured when assessing its relationship with farm size, as well as for drawing policy conclusions and recommendations.

The lack of an explicit focus on total factor productivity is a curious feature of the inverse relationship literature, especially given the early and widespread acknowledgement of its superiority over partial measures. As Barrett (1996) notes, this literature "habitually, perhaps cavalierly," uses physical yields and productivity synonymously. Total factor productivity is likely to be the measure most relevant to policy-makers formulating development strategies in poor economies where poverty alleviation is a pressing need and the productive use of all resources is paramount. In this light, we argue that the inverse relationship literature needs to shift its focus from land productivity to total factor productivity. In fact, empirical studies assessing the

productivity-farm size relationship in the developed world, such as Garcia et al. (1982), Alvarez and Arias (2004), and Rasmussen (2010), almost exclusively use measures of technical efficiency or total factor productivity.[2] Similarly, the literature estimating national level agricultural productivity is clear in its use of total factor productivity as a preferred measure (Fuglie, 2008; and Headey et al., 2010).

We illustrate the importance of choice of productivity measure with new empirical evidence on the farm size – productivity relationship in five regions of Brazil between 1985 and 2006. This period in Brazil provides an excellent case study because it includes regions with relatively advanced agricultural sectors, those characterized by more traditional agricultural production, and others experiencing rapid agricultural transformation, allowing us to assess the farm size – productivity relationship and its dynamics at different stages of agricultural development. Using a pseudo-panel of farms aggregated at the municipality and farm size levels, we show that estimating the farm size – productivity relationship using land productivity is potentially misleading. While we always identify an inverse relationship using land productivity, we find disparate results when using total factor productivity. In the modern agricultural regions of Brazil we find a direct relationship between farm size and total factor productivity, and in the rapidly transforming region of the Center-West we identify dynamics that suggest the inverse relationship is disappearing over time. The analysis highlights how the relationship between total factor productivity and farm size has evolved with modernization, shedding some light on the issues raised by Mill over 150 years ago.

---

[2] We have chosen not to focus on the literature estimating a stochastic production frontier to explore technical efficiency, as it is still an infrequent, albeit important, approach taken in the existing literature on developing countries.

The remainder of this paper is organized as follows. First, we seek to clarify the common measures, their relationships, and their advantages and limitations in empirical work. From this discussion, we lay out a framework for understanding how the farm size – productivity relationship may change over time, and how the choice of productivity measure matters for identifying those dynamics. We then present the empirical exercise, generating new evidence on the relationship between size and productivity in the macro regions of Brazil. Finally, we summarize and conclude with policy implications from the analysis.

II. Measures of Agricultural Productivity

Farm size may be related to a broad range of economic outcomes, such as efficiency, employment, poverty, inequality, food security, and growth. While these are important issues connected to the role of farm size in development, here, as with most of the literature on the inverse relationship (IR), we focus specifically on the concept of productivity. We do not attempt to address the issues of endogeneity or measurement error, nor do we attempt to explain the IR, as do many of the contributions in this field. Rather, we seek to clarify the relationships between the various productivity measures used, drawing conclusions on the importance of choice of measure when estimating the farm size – productivity relationship and generating insight into the impact that choice of measure may have on finding an IR.

The Unconditional Relationship between Land Productivity and Farm Size

Historically, land productivity is the most commonly used measure in the literature on the inverse relationship. Where alternative productivity measures are used, the relationship between gross land productivity and farm size is often a starting point because it is a benchmark

for the expansive existing literature on the IR. Land productivity, q, is a partial measure of productivity:

$$Land\ Productivity = \frac{Q}{A} = q = \psi_u(A) \tag{1}$$

where A is the area of the farm, Q is an index of agricultural output, q is agricultural output per unit of land, and $\psi_u(A)$ connotes that land productivity may be a function of farm size. In a world where farm size and land productivity are unrelated we have $\frac{\partial \psi_u(A)}{\partial A} = 0$. However, the regularity with which empirical work has found $\frac{\partial \psi_u(A)}{\partial A} < 0$ has led to the stylized fact that they are inversely related, and is precisely what has led to the abundant interest in the relationship and its potential explanations. Figure 1 displays this relationship using data from Brazil for the years 1985, 1996, and 2006. While the relationship is potentially non-linear and may not be monotonic, for now we focus on the first order approximation.

The relationship captured by $\psi_u(A)$ is unconditional (u) in the sense that it is the simple bivariate relationship between land productivity and farm size. Factors that may be causing or influencing this relationship have not been controlled for. Using land productivity as a measure is inherently limited—as would be any partial measure of productivity—whenever there is more than one factor of production. If other factors vary systematically with farm size then omitted variables may generate an IR due to more intensive input use by small farms, implying that a focus on the relationship between land productivity and the size of the farm may be misplaced. Similarly, analysis using different partial productivity measures may result in conflicting policy recommendations. Indeed, Sen's (1962) seminal contribution revealed precisely this type of systematic relationship between the intensity of labor use and farm size, leading to his formal

exposition of the dual labor market hypothesis (Sen 1966). Figure 2 illustrates the problem in the case of Brazil. While there is an inverse relationship between land productivity and farm size, there is a direct relationship between labor productivity and size. Analysis of the farm size and productivity relationship using labor productivity suggests that larger farms are more productive than are their smaller counterparts. Policy recommendations from the two partial measures of productivity would differ, underscoring the need for a comprehensive measure of productivity when identifying any relationship with farm size.

The Conditional Relationship between Land Productivity and Farm Size

It is curious that the unconditional relationship between land productivity and farm size continues to be used, even if in conjunction with more comprehensive measures of productivity. A more appropriate approach is to use a conditional relationship, where the relationship is conditioned on a vector of controls, $X(A)$, that may be correlated with both land productivity and farm size:

$$q = \psi_u(A) = g\big(X(A), \psi_c(A)\big) \tag{2}$$

The conditional (c) relationship, $\psi_c(A)$, should differ from the unconditional relationship to the extent that the conditioning controls explain the unconditional IR. For example, the impact of varying input intensities can be controlled for by including the inputs as controls, household heterogeneity can be controlled for with household fixed effects, market failures controlled for with regional fixed effects, and omitted variables such as soil quality can be introduced. This is a useful approach for exploring the theoretical channels that explain the IR and is a strategy commonly used by researchers in recent empirical studies of the farm size – productivity relationship (Assunção and Braido, 2007; and Barrett et al., 2010, among others).

As discussed above, partial measures such as land productivity are potentially misleading when there are other factors of production. At the very least, understanding any relationship between productivity and farm size requires empirical analysis that controls for the intensity with which other factors of production are used. A natural way to handle this is to include those factor intensities as conditioning variables. For exposition, assume that land, labor (L), and capital (K) are the only factors of production and that their intensities, labor per unit of land and capital per unit of land, are given by l and k, respectively. If these are the only controls then (2) becomes:

$$q = \psi_u(A) = g\big(k(A), l(A), \psi_c(A)\big) \tag{3}$$

From equation (3) we see that the IR as identified by the unconditional relationship between land productivity and farm size, $\frac{\partial \psi_u(A)}{\partial A}$, is composed of the relationship between capital intensity and farm size, labor intensity and farm size, and any conditional relationship between farm size and land productivity, $\frac{\partial \psi_c(A)}{\partial A}$. When differences in the use of other factors of production are controlled for, the conditional relationship between farm size and productivity captures a more comprehensive measure of productivity.

Exploring (3) highlights the ambiguity in how the land productivity and farm size relationship, as captured by $\psi_u(A)$, is related to the more general productivity and farm size relationship captured by $\psi_c(A)$. Differentiating (3) with respect to farm size shows:

$$\left(\frac{\partial \psi_u}{\partial A}\right) = \left(\frac{\partial g}{\partial k}\right)\left(\frac{\partial k}{\partial A}\right) + \left(\frac{\partial g}{\partial l}\right)\left(\frac{\partial l}{\partial A}\right) + \left(\frac{\partial g}{\partial \psi_c}\right)\left(\frac{\partial \psi_c}{\partial A}\right) \tag{4}$$

Assuming, quite reasonably, that output per unit of land is increasing in both capital and labor per unit of land, which is to say that $\frac{\partial g}{\partial k}$ and $\frac{\partial g}{\partial l}$ are both positive, it is plausible for the conditional relationship to be positive $\left(\frac{\partial \psi_c}{\partial A} > 0\right)$ even if the unconditional relationship is negative $\left(\frac{\partial \psi_u}{\partial A} < 0\right)$

if, as is often the case, $\frac{\partial k}{\partial A}, \frac{\partial l}{\partial A}$, or both are negative. In short, an unconditional IR is neither a necessary nor a sufficient condition for an inverse relationship between a broader measure of productivity and farm size as captured by $\psi_c(A)$.

When, as in (3), the conditional relationship includes all factors of production as controls, the approach is equivalent to estimating a production function. This conditional relationship can be interpreted as total factor productivity, arguably the most relevant productivity measure for policy. Most policy objectives – from poverty reduction to development and growth – are best informed by comprehensive measure of productivity that take into account the productivity with which all resources are jointly utilized. Estimating the farm size – productivity relationship by controlling for some, but not all, inputs correlated with farm size estimates a relationship with something less comprehensive than total factor productivity. Before returning to these issues, let us introduce an alternative and commonly used approach to measuring the IR, one that uses variations of profit as a measure of productivity.

Profit and Value-Added per Unit of Land

There are two variants of this approach. Start with the measure of profit:

$$\Pi = Q - p_L L - p_K K - p_A A \tag{5}$$

where p_L is the price of labor, p_K the price of capital, and p_A the price of land. In expression (5) the output quantity index, Q, is constructed using prices in the aggregation process, making it interpretable as the value of output. It is natural to think about the level of profit as the product of output and profit per unit of output: $Q\frac{\Pi}{Q}$. Regardless of whether the profit per unit of output rises or falls with size, we would expect the level of output to dominate in the determination of

the level of profit. A large farm that produces 1,000 units of output should generate more profit than a small farm that produces 10 units. The level of profit, then, is not a particularly good measure for comparing the productivity of farms of different sizes.

It is not profits per se that matter, but rather profitability. This requires the transformation of the profit level into a profit rate. Profit per unit of land, as used by Carletto et al. (2013), is one approach:

$$Profit\ per\ unit\ of\ land = \pi_A = \frac{\Pi}{A} = q - p_L l - p_K k - p_A = q - c = \phi(A) \qquad (6)$$

where c is the cost per unit of land incurred in producing q. Profit per unit of land is a productivity measure that controls for the levels of other inputs additively, providing an improvement over land productivity. Measures of value added are similar, however they fall short of profit measures as they control only for intermediate inputs and not the complete set of factors of production. Despite being an improvement over land productivity and value-added, profit per unit of land is itself problematic because it is fundamentally a partial measure. The finding of a systematic inverse relationship with farm size, $\frac{\partial \phi(A)}{\partial A} < 0$, is of limited use because this partial measure does not control for the profitability with which other inputs are used:

$$\pi_A = \frac{\Pi}{A} = \frac{\Pi}{K}\frac{K}{A} = \pi_K * k \qquad (7)$$

Here we see that the profit per unit of land can be rewritten as the product of profit per unit of capital, π_K, and capital intensity. An observed inverse relationship between profit per unit of land and farm size could be associated with declining capital intensity as farm size increases, even if profit per unit of capital is increasing. If this were true, then the use of one partial measure or the other would lead to conflicting policy recommendations, and again a more comprehensive measure is preferred.

TFP and Profit Rate

Comprehensive measures, of either productivity or profitability, are the appropriate means to measure the efficiency of resource use. Berry and Cline (1979), along with Binswanger et al. (1995), have argued that total factor productivity (TFP) or the profit rate are the preferable and correct choices for productivity measures when discussing farm size. We concur. Total factor productivity is a comprehensive productivity measure defined as the ratio of output to all inputs used. TFP can be written as:

$$TFP = \frac{Q}{Input\ Quantity\ Index} = \varphi(A) \tag{8}$$

TFP effectively captures the productivity with which all inputs are used in the production process. It measures units of output for every unit of input used, regardless of its type, and in this sense is a comprehensive measure of productivity. If this measure is a function of farm size, i.e. $\frac{\partial \varphi(A)}{\partial A} \neq 0$, then there is an unambiguous difference in how productively farms of different sizes utilize resources in agricultural production. An understanding of the determinants of $\varphi(A)$ would support effective policy design, whether the objective is poverty reduction or economic development, because these are concerned with the use of all resources available to farms. Although this is widely acknowledged, an explicit focus on TFP is seldom the approach of empirical analyses of the IR in developing economies.

TFP can be written as a function of a profit rate, where profit is normalized by the total value of inputs used in production. If land, labor, and capital are the only three factors of production, then:

$$TFP = \frac{\Pi}{p_L L + p_K K + p_A A} + 1 \tag{9}$$

This expression requires that the input quantity index used in the calculation of TFP is constructed using factor prices as weights. The profit rate described in (9) is the rate most meaningful for assessing agricultural productivity in that it is consistent with TFP. Whereas TFP looks only at the gross value of production, the profit rate considers the net value of production. Any IR found using this profit rate should also be found using TFP as the productivity measure. Whereas profit is linear in costs, TFP controls for costs multiplicatively and is log-linear.

An alternative interpretation of the profit rate results from normalization by the total value of assets used in agricultural production. Consider, for instance, a potential farmer with financial capital looking to invest in farming. If there are competitive factor markets then the value of the capital used to rent land, physical capital, and labor to run the operation is identically equal to the costs of production. Binswanger et al. (1995) advocate normalizing profit by "capital invested" or "assets." This is appropriate as long as measured assets are restricted to those used in production and do not include all of household wealth (as is sometimes done in the literature). If other forms of wealth are erroneously included, and farm size and wealth are positively related, then an IR may appear to emerge even when profit per unit of invested capital is constant or increasing with farm size.

The use of such a profit rate is likely to be of limited practical applicability in the context of developing country agriculture because of the analyst's need to observe input prices generated by well-functioning input markets. In practice, this is often unrealistic, requiring assumptions on input prices that potentially bias the estimates. Berry and Cline (1979) offer such an example, where assumptions on the value of household labor alter the observed farm size – TFP relationship.

Linking TFP and Land Productivity

Consider first the relationship between TFP and the conditional relationship captured when analyzing a production function. Assume that the production function is homogenous of degree t so that $f(\lambda L, \lambda K, \lambda A) = \lambda^t f(L, K, A)$. Constant returns to scale (CRS) holds if $t = 1$, while decreasing returns (DRS) are observed if $t < 1$ and increasing returns (IRS) if $t > 1$. Setting $\lambda = \frac{1}{A}$ conveniently implies that the production function can be written as:

$$f(l, k, 1) = A^{-t} f(L, K, A) \tag{10}$$

and dividing output by farm size gives us land productivity:

$$q = \frac{f(L,K,A)}{A} = A^{t-1} f(l, k, 1) \tag{11}$$

As an example, assume a standard Cobb-Douglas production function where *T* is the unobserved measure of total factor productivity and production is a function of labor, capital and land:

$$f(L, K, A) = T L^\alpha K^\beta A^\gamma \tag{12}$$

If, as has often been confirmed, CRS holds, then farm size disappears from the right hand side of (12) after dividing through by farm size. If not, then the natural log of the production function takes the form:

$$lnq = (t - 1)lnA + lnT + \alpha lnl + \beta lnk \tag{13}$$

If, as in (8), there exists a relationship between total factor productivity and size, $\varphi(A)$, we have:

$$lnq = (t - 1)lnA + ln\varphi(A) + \alpha lnl + \beta lnk \tag{14}$$

From (14) it is clear that the conditional relationship identified in (3), $\psi_c(A)$, is composed of the relationship between TFP and farm size as well as any deviations from CRS in the production function.

Equation (14) highlights three useful features of the production function approach. First, if CRS holds then the conditional relationship, $\psi_c(A)$, captures the relationship between TFP and farm size, $\varphi(A)$. Second, if CRS does not hold then it will be difficult to empirically differentiate whether a conditional relationship is driven by non-CRS, a relationship between TFP and farm size, or a combination of the two. This highlights the importance of explicitly investigating returns to scale and, if this is not done, to interpret any observed conditional relationship as potentially coming from non-CRS. Third, if the addition of any controls other than the factor intensities explain differences in the conditional relationship, which is composed of non-CRS and the TFP-farm size relationship, they should be interpreted as such.

Thus, when empirical researchers estimate a production function to explore the relationship between farm size and productivity they are, in effect, estimating the relationship between farm size and TFP, not farm size and land productivity. All too often, empirical work that takes this approach shows the unconditional relationship first (often non-parametrically), followed by a series of regressions of production functions that include other controls, followed by interpretation of the two approaches as if they were exploring the same relationship. However, the conditional and unconditional relationships are by no means the same, can plausibly take different signs, and will almost certainly have different magnitudes.

The relationship between TFP and output per unit of land can be explored further. TFP is a unit-less measure, but multiplying and dividing by $\frac{1}{A}$ allows the measure to be rewritten as:

$$TFP = \frac{q}{inputs\ per\ unit\ of\ land} = \frac{q}{\tau} = \varphi(A) \qquad (15)$$

From the expression above, TFP can be thought of as land productivity normalized by input costs per unit of land, τ.

Taking a derivative with respect to farm size:

$$\frac{\partial TFP}{\partial A} = \left(\frac{\partial q}{\partial A}\right)\left(\frac{1}{\tau}\right) - \frac{\left(\frac{\partial \tau}{\partial A}\right)q}{\tau^2} \tag{16}$$

Employing a little bit of algebra (see Appendix A):

$$\frac{\partial TFP}{\partial A} = \left(\frac{\partial q}{\partial A}\right)\left(\frac{1}{\tau}\right)\left[\frac{\varepsilon_{q,A} - \varepsilon_{\tau,A}}{\varepsilon_{q,A}}\right] \tag{17}$$

where $\varepsilon_{q,A}$ is the elasticity of land productivity with respect to farm size, and $\varepsilon_{\tau,A}$ is the elasticity of input use per unit of area with respect to farm size. If there is an empirically observed inverse relationship between the partial measure and farm size such that $\frac{\partial q}{\partial A} < 0$, then we know $\varepsilon_{q,A}$ is negative. This implies that one of two possibilities must hold:

(i) $\qquad \frac{\partial TFP}{\partial A} < 0$ and $\varepsilon_{q,A} < \varepsilon_{\tau,A}$

(ii) $\qquad \frac{\partial TFP}{\partial A} > 0$ and $\varepsilon_{q,A} > \varepsilon_{\tau,A}$

If (i) is true then an IR between a partial measure and farm size reflects an IR between productivity and farm size as measured by TFP. When this is the case then either input use per unit of land is increasing in farm size or it is decreasing, but slower than the rate at which output per unit of land is decreasing. If (ii) is true then use of a partial measure is generating an incorrect indication about the productivity and farm size relationship, and TFP is actually directly related to farm size. However, this requires that $0 > \varepsilon_{q,A} > \varepsilon_{\tau,A}$. In such a case, input use per unit of land is negatively related to farm size and is relatively elastic compared to output per unit of land. This discussion highlights the conclusion that an IR between a partial measure of productivity and farm size is neither necessary nor sufficient for the existence of an IR between farm size and a comprehensive measure of productivity such as TFP.

Discussion of the Dynamics of TFP and Partial Measures

The conditions set out in (i) and (ii) provide a framework for considering how a modernizing agricultural sector can lead to a changing farm size – productivity relationship. Depending upon the stage of development and the institutional structure, partial measures of productivity may fail to capture the dynamics of the farm size – productivity relationship.

Consider first an economy at an early stage of development, where a traditional agricultural sector relies predominantly on land and labor as inputs to production and land, labor, and credit markets exhibit imperfections. If an IR between land productivity and farm size exists in such a world, the underlying mechanisms are likely to be Sen's dual labor market hypothesis and the monitoring costs associated with larger farms hiring wage labor. Larger farms must incur supervision costs to maintain a given labor intensity, implying that, even if labor intensity falls with farm size, $\varepsilon_{\tau,A}$ may be positive and (i) holds. Even if labor intensity is falling with farm size fast enough such that $\varepsilon_{\tau,A}$ is negative, the existence of supervision costs implies that $\varepsilon_{\tau,A}$ is likely to be less elastic than $\varepsilon_{q,A}$. Again, condition (i) holds and TFP, as with land productivity, will have an inverse relationship with farm size. In such an economy land productivity may provide an adequate proxy for TFP, even if the magnitudes of the two relationships differ.

In a second stage of agricultural development, advances in agricultural technology lead to mechanization, making capital a more important class of inputs for those farms that are able to adopt modern technologies. Such a stage may continue to be characterized by struggling institutions and imperfections in credit and labor markets, and large farms may more readily be able to adopt new technology and substitute away from labor. Smaller farms may continue to employ labor more intensively in the presence of labor market imperfections, implying the

possibility that an inverse relationship between land productivity and farm size continues to exist even as improved access to capital for larger farms attenuates the relationship. The substitution away from labor towards capital avoids costly monitoring, moving large farms towards a more efficient mix of factors of production. In this context, condition (ii) might hold, with input use per unit of land ($\varepsilon_{\tau,A}$) falling more quickly than output per unit of land ($\varepsilon_{q,A}$). A direct relationship between TFP and farm size could emerge, even as an IR continues to exist for land productivity. Sadoulet and De Janvry (1995) discuss such a case, where a move to capital intensive agricultural production could generate a positive relationship between farm size and TFP. Unlike in the initial stage of development, mechanization and an increasingly important role for capital have rendered land productivity a misleading indicator of the relationship between farm size and TFP. This is consistent with Carter's (1984) findings.

In a developed economy, a third stage of agricultural development may realign the relationships between TFP, land productivity and farm size. Institutions improve as economic development continues, and distortions in the labor and capital markets begin to disappear. While capital-intensive agriculture becomes widespread, sources of economies of scale other than mechanization—such as those that derive from information technologies—begin to come into play. Larger farms also have better access to managerial and technical expertise, and they have advantages in marketing and distribution. In such an environment, the inverse relationship between land productivity and farm size could disappear, implying $\varepsilon_{q,A} \geq 0$. Both land productivity and TFP could conceivably exhibit a direct relationship with farm size.

The discussion above suggests several important implications for the IR literature. First, it highlights how partial measures of productivity are unlikely to serve as appropriate proxies for

comprehensive productivity measures such as TFP in the context of modernizing agricultural

sectors. Second, we should expect the farm size – productivity relationship to be dynamic,

evolving with the economy. Third, a focus on the choice of productivity measure and the

dynamics involved help to reconcile disparate empirical results. Empirical findings of an inverse

farm size – land productivity relationship in economies at early and medium stages of

development are fully consistent with empirical findings of a direct farm size – TFP relationship

in more advanced economies. Thus, while a growing body of within-country studies has identified

a diminishing inverse relationship over time (Deininger et al., 2018), or non-linearities in the farm

size – productivity relationships (Foster and Rosenzweig, 2017; Muyanga and Jayne, 2016; Kimhi,

2006; Helfand and Levine, 2004), the multi-country study led by Rada and Fuglie (2018) provides

evidence for a U-shape across countries. These findings can be understood in the framework

presented above.

III. Empirical Analysis

We now provide an illustrative example using data on Brazilian agriculture. The intention

here is not to explain the relationship between farm size and productivity by controlling for its

potential determinants. Rather, we seek to use a regional analysis within Brazil to highlight how

the choice of measure influences the observed relationship, potential non-linearities, and how

the patterns change across stages of agricultural development.

Data and Variables

The data come from the 1985, 1995/1996, and 2006 rounds of the Brazilian agricultural

census. For confidentiality reasons, we construct a pseudo panel in which all farms in the census

are aggregated into five farm size classes for each municipality in Brazil.[3] The changing composition of Brazil's municipalities requires the construction of minimum comparable areas (AMCs) to generate a dataset of geographic units that is consistent over time. This process aggregates the 5,175,636 farms across the 5,548 Brazilian municipalities in 2006 into 3,861 AMCs. For each farm size class in each AMC we generate a "representative farm" that characterizes the behavior of farms of that class in that AMC, effectively assuming homogeneity among farms within each AMC farm size class.

We begin with 47,365 representative farms for all of Brazil across the three survey years. Due to concern about the comparability of a small number (84) of extremely large farms, we remove all representative farms in the Northeast and South over 4,000 ha and all of those over 5,000 ha in the North, the Southeast, and the Center-West. We identify land productivity outliers in two stages. First, we cut the top and bottom 1% of the land productivity distribution from the sample. Second, using a quadratic specification we regress land productivity on farm size with AMC fixed effects and survey year dummy variables. From this regression we identify all representative farms with land productivity greater than four standard deviations from the conditional mean as outliers. Together, the data cleaning exercises remove 2.3% of the initial sample.[4]

---

[3] The size classes are 0-5 ha, 5-20 ha, 20-100 ha, 100-500 ha, and 500+ ha. To protect the confidentiality of the farms, the Brazilian Institute of Geography and Statistics (IBGE) requires that each aggregate observation have at least 3 farms. As the aggregation was conducted on site prior to analysis, we are not able to expand the number of farm size bins. However, previous work using the underlying Brazilian census data found little difference across alternative bin specifications (Helfand et al., 2014; Moreira et al., 2007; Helfand and Levine, 2004).

[4] See Table A1 of Appendix B for the results of data cleaning from each stage of the process. Sensitivity analyses using alternative restrictions on the data had no qualitative impact on our results.

Output is measured as the real value (R$2006) of total production, valued using regional average prices for the given year and deflated with a price index developed from the data of Gasques et al. (2010). Farm size is measured in hectares (ha). The data is self-reported and, while measurement error is potentially an issue for identifying the farm size – productivity relationship, the aggregation of the data can help to mitigate the problem. The literature is inconclusive regarding the potential effects of measurement error – whereas Desiere and Jolliffe (2017) find that measurement error may fully explain the IR, Lamb (2003) and Dillon et al. (2016) conclude that measurement error explains only part of the IR, and Carletto et al. (2013) finds evidence that, to the contrary, controlling for measurement error can strengthen the IR. In a recent paper by Abay et al. (2018), the authors suggest that in the presence of multiple variables that exhibit measurement error—such as output and farm size—there is no guarantee that correcting only one will reduce the bias of the estimated coefficients.

Additional factors of production include quantity indices for family labor, purchased inputs, and capital. Counts of male, female, and child family members working on each farm are used to develop a family labor index measured in adult male equivalents. The real value (R$2006) of purchased inputs, including expenditure on fertilizer, seeds, hired labor, fuel, energy, soil amendments, and other inputs, are calculated with the price index used for output. A measure of the total capital stock is calculated as a quantity index comprised of machine, animal, and tree capital stock sub-indices following Moreira et al. (2007) and Butzer et al. (2012). The machine capital stock index values tractors, trucks, and other agricultural equipment using a constant set of sale prices drawn from the Instituto de Economia Agrícola in São Paulo. The stock of animal capital is measured in cattle equivalents of nine distinct animal stocks and aggregated with a set

of time-invariant relative prices. The stock of tree capital is measured as the present discounted value of expected future profits for thirteen different tree crops, using region-specific estimates of expected profits. The sub-indices are aggregated using region-specific weights estimated by regressing output on the capital stock sub-indices in 1985. Further discussion of the dataset can be found in Rada et al. (2018).

By focusing on a regional analysis we are able to examine the relationship between farm size and productivity in light of each region's characteristics and stage of development. The five macro-regions of Brazil differ in both the type of predominant agricultural activities and the degree of modernization. They include the Amazon rainforest in the North, a large semi-arid region in the Northeast, a highly mechanized and commercial agriculture in the Southeast, a predominance of family farms in the South, and the Cerrado (savannahs) of the Center-West where grains have rapidly expanded and agriculture has modernized in recent decades. We restrict attention to the North, Center-West, and Southeast, three macro-regions that capture sufficient regional variation in Brazilian agriculture to illustrate our argument. Descriptive statistics for output and input intensities for these regions are shown in Table A2 of the appendix. Differences in input intensities reflect the heterogeneity in agricultural production across regions. The more traditional agricultural region in the North relies more heavily on family labor, whereas the mechanized Southeast and Center-West use capital and purchased inputs more intensively.

Empirical Methodology

We estimate an average production function assuming a Cobb-Douglas technology. Output and inputs for a farm in AMC m of size s in year t are normalized by area, A_{mst}. Estimating the model using intensities forces any deviation from CRS into the estimated relationship

21

between farm size and productivity. Survey year specific dummy variables for five farm size classes, δ_{st}, are used to flexibly capture the relationship between farm size and TFP. The farm size class 0-5 ha in 1985 is excluded and used as a reference. While this structure allows the farm size and productivity relationship to change over time, the technology coefficients are assumed to be time invariant. This assumption forces technical change into our measure of TFP. The estimated equation takes the form:

$$lny_{mst} = \beta_0 + \beta_1 lnk_{mst} + \beta_2 lnl_{mst} + \beta_3 lnx_{mst} + \delta_{st} + \lambda_m + \varepsilon_{mst} \qquad (18)$$

where y_{mst} is aggregate output per unit of land and k_{mst}, l_{mst}, and x_{mst} are the factors of production per unit of land (capital, family labor, and purchased inputs including hired labor, respectively), and λ_m are AMC fixed effects. The parameters are estimated using ordinary least squares, with standard errors clustered at the AMC level. Because the number of farms represented by each representative farm varies, each observation is weighted by the number of farms that it represents. With the above approach the TFP of a farm size bin in a given period is calculated by:

$$TFP_{st} = e^{\beta_0 + \delta_{st}} \qquad (19)$$

A TFP index is then calculated for each size class in each period using the size class 0-5 ha in 1985 as a base level set to 100.

Empirical Results

Figure 3 shows the unconditional relationship between land productivity and farm size class for the three regions under study. Despite considerable regional heterogeneity in their agricultural activities and agrarian structures, each region mirrors the country as a whole in

displaying a strong inverse relationship between land productivity and farm size. If anything, the relationship appears to be growing stronger over time.

The estimated coefficients from region-specific estimates of equation (18) are shown in Table A3 of the appendix, which generate the TFP estimates presented in Figures 4 through 6. In the North (Figure 4), we estimate an inverse relationship between farm size and TFP. It is not, however, a linear relationship, but rather an emerging U-shaped inverse relationship with farms over 500 ha becoming more productive than medium-sized farms. The significance tests in Table 1 confirm this, showing that while the productivity of farms between 20 ha and 500 ha is statistically less than the smallest farms in all periods, the largest farms are not statistically different than the smallest farms. Thus, while a strong negative relationship would be found in this region when using land productivity, a U-shaped relationship begins to emerge when TFP is used and linearity is not imposed.

The Center-West (Figure 5) demonstrates a more dynamic pattern. Table 1 shows that the farm size – TFP relationship in the Center-West in 1985 looked very similar to the inverse relationship in the North. However, by 2006 the inverse relationship had disappeared in the Center-West, with the TFP of all farm sizes being statistically indistinguishable from that of the smallest farms. The point estimates show that the largest farms in the Center-West were 45% less productive than the smallest farms in 1985, yet by 2006 they were 14% more productive, albeit statistically insignificantly so. Once again, a U-shape emerges, driven by rapid growth of the productivity of larger farms. This is the clearest case of a strong inverse relationship becoming reversed over the 21 year period. Using land productivity to measure the farm size – productivity

relationship in a rapidly modernizing agricultural region such as the Center-West would completely miss this transformation.

The Southeast, in contrast, shows a positive non-linear relationship between farm size and TFP. The relationship was statistically flat in 1985, although the point estimates show that even in 1985 the largest farms were 26% more productive than the smallest. Rapidly rising TFP at the upper end of the farm size distribution makes the relationship more positive over time, and by 2006 the largest farms were 49% more productive than the smallest, and statistically so. Once again, the relationship appears non-linear. This contrasts sharply to the persistent IR found in the Southeast when using land productivity as a measure.

In comparison to much of the development literature surrounding the IR, the Brazilian data used here represent a very heterogeneous group of farms and span a much greater range of farm sizes. A more accurate comparison group to the international literature might be farms less than 100 ha, which indeed make up approximately 90% of all Brazilian farms. Even when restricting our analysis to this subset of farms, the use of land productivity would still show a marked inverse relationship while the use of TFP would reveal a negative relationship that has disappeared in the more modernizing regions. Perhaps more importantly, inclusion of the largest farm size class reveals that these farms have notably higher productivity in the more modern regions, and it is only when TFP is used that this becomes apparent. These are commercial farms that are unlikely to be included in most household surveys in developing countries.[5]

---

[5] The empirical results obtained here are comparable to those reported at the national level by Rada et al. (2018) using a similar dataset. One difference is that they find somewhat faster TFP growth for the smallest farm size class, resulting in a more pronounced U-shape in 2006. The principal differences in empirical methodology are that they estimate TFP growth separately for each farm size class, and do not explore regional heterogeneity.

IV. Conclusions and Policy Implications

We have sought to address an important weakness of the development economics literature on the inverse relationship between farm size and productivity. We argued that a variety of productivity measures are used when estimating this relationship, that the choice of measure matters for the identification and interpretation of the farm size – productivity relationship, and that total factor productivity is, in most cases, the preferred and most appropriate measure. Furthermore, we argued that a commonly used measure—land productivity—is problematic and potentially misleading when used in modernizing agricultural contexts or when assessing a full range of farm sizes. While this point has been noted in the literature, the use of land productivity and other partial measures of productivity remains widespread. Furthermore, where movements are made towards use of a comprehensive measure, such as TFP, the different measures are often used synonymously and there is a lack of clarity in their interpretation. Our conceptual discussion provides a framework for assessing the implications of the choice of productivity measure. Theoretically, it is clear that an inverse relationship between land productivity and farm size is neither necessary nor sufficient for an inverse relationship to exist between farm size and TFP.

How much does this critique matter? We conduct an empirical analysis of three regions in Brazil using a pseudo-panel from 1985 to 2006 to contrast the land productivity – farm size relationship with the TFP – farm size relationship. In short, the choice of productivity measure matters greatly. As in many developing country contexts, there exists an inverse relationship between land productivity and farm size for Brazil, and within each of its macro-regions in every period. In contrast, the TFP and farm size relationship varies across time and space. The regional

analysis of the TFP and farm size relationship shows 1) land productivity is not always an appropriate proxy for TFP; 2) the relationship is dynamic, changing with agricultural modernization; 3) the relationship is non-linear, often characterized by a U-shape; and 4) the very largest farms, such as those with more than 500 ha, are important to consider when assessing any relationship between farm size and productivity.

From a policy perspective, our findings have important implications for the debate about the future of small farms in developing countries. When using TFP, we see that superior productivity of small farms in traditional agricultural contexts is fully consistent with emergent productivity advantages for larger commercial farms in modernizing agricultural sectors. As economies develop, superior productivity may not continue to provide a valid argument for the importance and future of small farms, as we expect larger farms to play a more important role in driving agricultural productivity growth. As such, it is increasingly unlikely that redistributive land reform could positively impact both equity and efficiency. This does not, however, imply that small farms will, nor should they, disappear. We expect them to remain important for generating livelihoods for rural households, providing food security, and contributing to the development of rural economies. Productivity gains among small farmers will also continue to be essential for poverty alleviation. Thus, policies that target productivity gains for small farmers, such as those that build human capital, facilitate adoption of new technologies, and enhance access to markets via a reduction in transactions costs, will continue to be indispensable for reducing rural poverty in developing countries.

References

Abay, Kibrom A., Gashaw T. Abate, Christopher B. Barrett, Tanguy Bernard, "Correlated Non-Classical Measurement Errors, 'Second Best' Policy Inference and the Inverse Size-Productivity Relationship in Agriculture." Draft, February 2018, 41 pages.

Ali, Daniel A., and Klaus Deininger. "Is There a Farm Size–Productivity Relationship in African Agriculture? Evidence from Rwanda." *Land Economics* 91.2 (2015): 317-343.

Alvarez, Antonio, and Carlos Arias. "Technical efficiency and farm size: a conditional analysis." *Agricultural Economics* 30, no. 3 (2004): 241-250.

Assunção, Juliano J., and Luis HB Braido. "Testing household-specific explanations for the inverse productivity relationship." *American journal of agricultural economics* 89, no. 4 (2007): 980-990.

Bardhan, Pranab K. "Size, productivity, and returns to scale: An analysis of farm-level data in Indian agriculture." *Journal of political Economy* 81, no. 6 (1973): 1370-1386.

Barrett, Christopher B. "On price risk and the inverse farm size-productivity relationship." *Journal of Development Economics* 51, no. 2 (1996): 193-215.

Barrett, Christopher B., Marc F. Bellemare, and Janet Y. Hou. "Reconsidering conventional explanations of the inverse productivity–size relationship." *World Development* 38, no. 1 (2010): 88-97.

Benjamin, Dwayne. "Can unobserved land quality explain the inverse productivity relationship?." *Journal of Development Economics* 46, no. 1 (1995): 51-84.

Berry, R. Albert, and William R. Cline. *Agrarian structure and productivity in developing countries: a study prepared for the International Labour Office within the framework of the World Employment Programme*. Johns Hopkins Univ. Press, (1979).

Bevis, L., and Christopher B. Barrett. *Close to the edge: High productivity at plot peripheries and the inverse size-productivity relationship*. Working paper, 2017.

Bhalla, Surjit S., and Prannoy Roy. "Mis-specification in farm productivity analysis: the role of land quality." *Oxford Economic Papers* 40, no. 1 (1988): 55-73.

Binswanger, Hans P., Klaus Deininger, and Gershon Feder. "Power, distortions, revolt and reform in agricultural land relations." *Handbook of development economics* 3 (1995): 2659-2772.

Butzer, R., Y. Mundlak, and D. Larson. "Measures of fixed capital in agriculture." Chapter 15 in Fuglie, K., S. Wang, and V. Ball (Eds.), *Productivity Growth in Agriculture: An International Perspective*. CAB International, Cambridge, MA, (2012).

Carletto, Calogero, Sara Savastano, and Alberto Zezza. "Fact or artifact: The impact of measurement errors on the farm size–productivity relationship." *Journal of Development Economics* 103 (2013): 254-261.

Carter, Michael R. "Identification of the inverse relationship between farm size and productivity: an empirical analysis of peasant agricultural production." *Oxford Economic Papers* 36, no. 1 (1984): 131-145.

Deolalikar, Anil B. "The inverse relationship between productivity and farm size: a test using regional data from India." *American Journal of Agricultural Economics* 63, no. 2 (1981): 275-279.

Deininger, Klaus, Songqing Jin, Yanyan Liu, and Sudhir K. Singh "Can Labor-Market Imperfections Explain Changes in the Inverse Farm Size–Productivity Relationship? Longitudinal Evidence from Rural India." Land Economics 94.2 (2018): 239-258.

Desiere, Sam, and Dean Jolliffe. "Land productivity and plot size: Is measurement error driving the inverse relationship?." *Journal of Development Economics* 130 (2018): 84-98.

Dillon, Andrew, Sydney Gourlay, Kevin McGee, and Gbemisola Oseni. *Land measurement bias and its empirical implications: evidence from a validation exercise*. The World Bank, (2016).

Eastwood, Robert, Michael Lipton, and Andrew Newell. "Farm size." *Handbook of Agricultural Economics* 4 (2010): 3323-3397.

Eswaran, Mukesh, and Ashok Kotwal. "Access to capital and agrarian production organisation." *The Economic Journal* 96, no. 382 (1986): 482-498.

Feder, Gershon. "The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints." *Journal of development economics* 18, no. 2-3 (1985): 297-313.

Foster, Andrew D., and Mark R. Rosenzweig. *Are there too many farms in the world? Labor-market transaction costs, machine capacities and optimal farm size*. No. w23909. National Bureau of Economic Research, (2017).

Fuglie, Keith O. "Is a slowdown in agricultural productivity growth contributing to the rise in commodity prices?" *Agricultural Economics* 39 (2008): 431-441.

Garcia, Philip, Steven T. Sonka, and Man Sik Yoo. "Farm size, tenure, and economic efficiency in a sample of Illinois grain farms." *American Journal of Agricultural Economics* 64, no. 1 (1982): 119-123.

Gasques, José Garcia, Eliana Teles Bastos, Mirian Rumenos Piedade BACCHI, and Constanza Valdes. "Produtividade total dos fatores e transformações da agricultura brasileira: análise dos dados dos censos agropecuários." *A agricultura brasileira: desempenho, desafios e perspectivas. Brasília: Ipea* (2010): 19-44.

Gourlay, Sydney, Talip Kilic, and David Lobell. *Could the debate be over? errors in farmer-reported production and their implications for the inverse scale-productivity relationship in Uganda*. The World Bank, (2017).

Headey, Derek, Mohammad Alauddin, and DS Prasada Rao. "Explaining agricultural productivity growth: an international perspective." *Agricultural Economics* 41, no. 1 (2010): 1-14.

Helfand, Steven M., Ajax R.B. Moreira and Edward W. Bresnyan Jr. "Agricultura Familiar, Produtividade e Pobreza no Brasil: Evidências do Censo Agropecuário 2006." Chapter 11 in Schneider, Sergio, Brancolina Ferreira, and Fabio Alves (eds.), *Aspectos Multi-Dimensionais da Agricultura Brasileira: Diferentes Visões do Censo Agropecuário 2006.* Brasília: IPEA (2014): 279-311.

Helfand, Steven M., and Edward S. Levine. "Farm size and the determinants of productive efficiency in the Brazilian Center-West." *Agricultural economics* 31, no. 2-3 (2004): 241-249.

Heltberg, Rasmus. "Rural market imperfections and the farm size—productivity relationship: Evidence from Pakistan." *World Development* 26, no. 10 (1998): 1807-1826.

Henderson, Heath. "Considering Technical and Allocative Efficiency in the Inverse Farm Size–Productivity Relationship." *Journal of Agricultural Economics*, 66, no. 2 (2015): 442–469.

Kagin, Justin, J. Edward Taylor, and Antonio Yúnez-Naude. "Inverse productivity or inverse efficiency? Evidence from Mexico." *The Journal of Development Studies* 52, no. 3 (2016): 396-411.

Key, Nigel, William McBride, and Roberto Mosheim. "Decomposition of total factor productivity change in the US hog industry." *Journal of Agricultural and Applied Economics* 40, no. 1 (2008): 137-149.

Kimhi, Ayal. "Plot size and maize productivity in Zambia: is there an inverse relationship?." *Agricultural Economics* 35, no. 1 (2006): 1-9.

Lamb, Russell L. "Inverse productivity: Land quality, labor markets, and measurement error." *Journal of Development Economics* 71, no. 1 (2003): 71-95.

Larson, Donald F., Keijiro Otsuka, Tomoya Matsumoto, and Talip Kilic. "Should African rural development strategies depend on smallholder farms? An exploration of the inverse-productivity hypothesis." *Agricultural Economics* 45, no. 3 (2014): 355-367.

Lipton, Michael. *Land reform in developing countries: Property rights and property wrongs*. Routledge, (2009).

Moreira, Ajax R. B., Steven M. Helfand, and Adriano M. R. Figueiredo. "Explicando as Diferenças na Produtividade Agrícola no Brasil." Texto para Discussão 1254. IPEA (2007).

Muyanga, Milu, and T. S. Jayne. "Is small still beautiful? The farm size-productivity relationship revisited." In *Conference on Land Policy in Africa*. (2014).

Rada, Nicholas E., and Keith O. Fuglie. "New perspectives on farm size and productivity." *Food Policy* (2018).

Rada, Nicholas E., Steven M. Helfand, and Marcelo M. Magalhães. "Agricultural productivity growth in Brazil: Large and small farms excel." *Food Policy* (2018).

Rosenzweig, Mark R., and Hans P. Binswanger. "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments." *The Economic Journal* 103, no. 416 (1993): 56-78.

Sadoulet, Elisabeth, and Alain De Janvry. *Quantitative development policy analysis*. Vol. 1. Baltimore: Johns Hopkins University Press, (1995).

Sen, Amartya K. "An aspect of Indian agriculture." *Economic Weekly* 14, no. 4-6 (1962): 243-246.Sen, Amartya K. "Peasants and Dualism with or without Surplus Labor." *Journal of political Economy* 74, no. 5 (1966): 425-450.

Figure 1: Farm Size and Land Productivity, Brazil (logs)
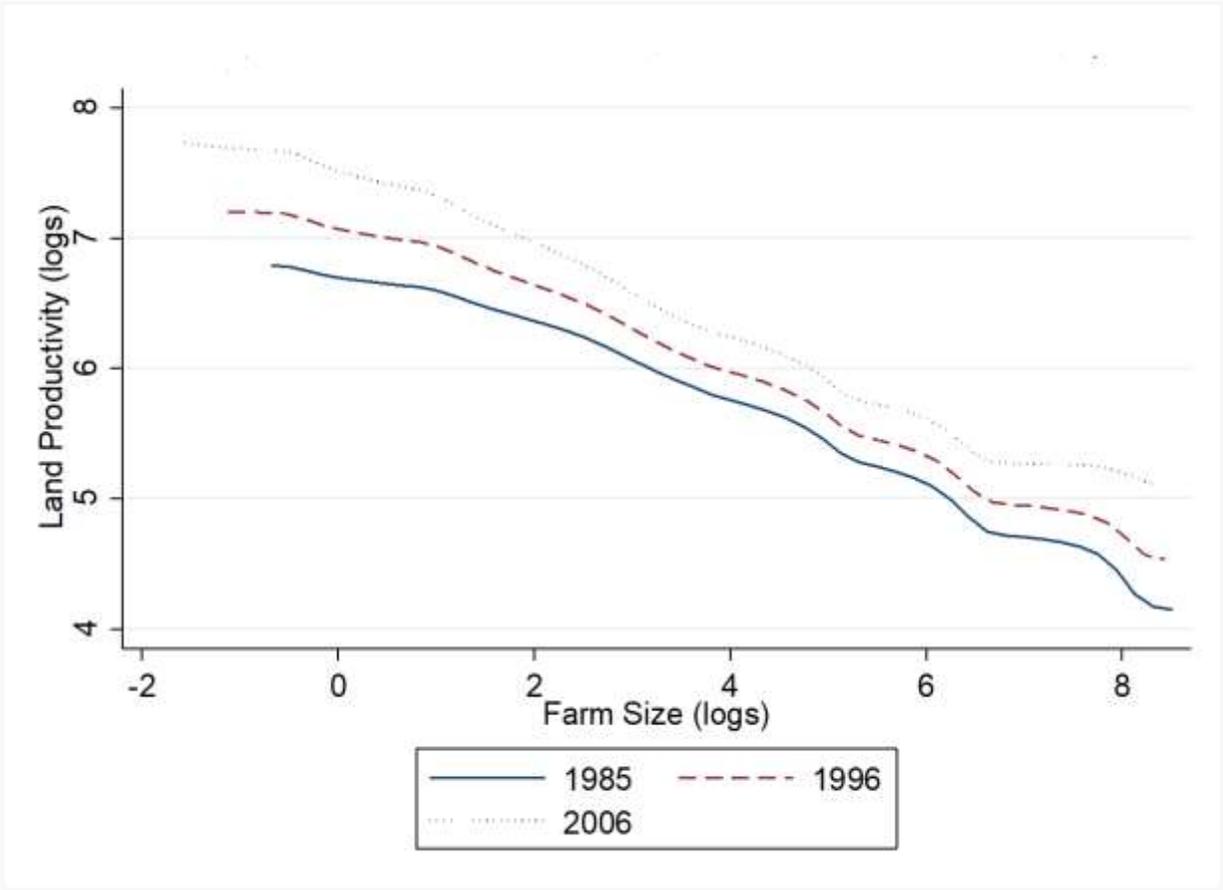
Note: Smoothed as a local polynomial regression with bandwidth of 1.25 and Epanechnikov kernel.

Figure 2: Land and Labor Productivity, Brazil 2006
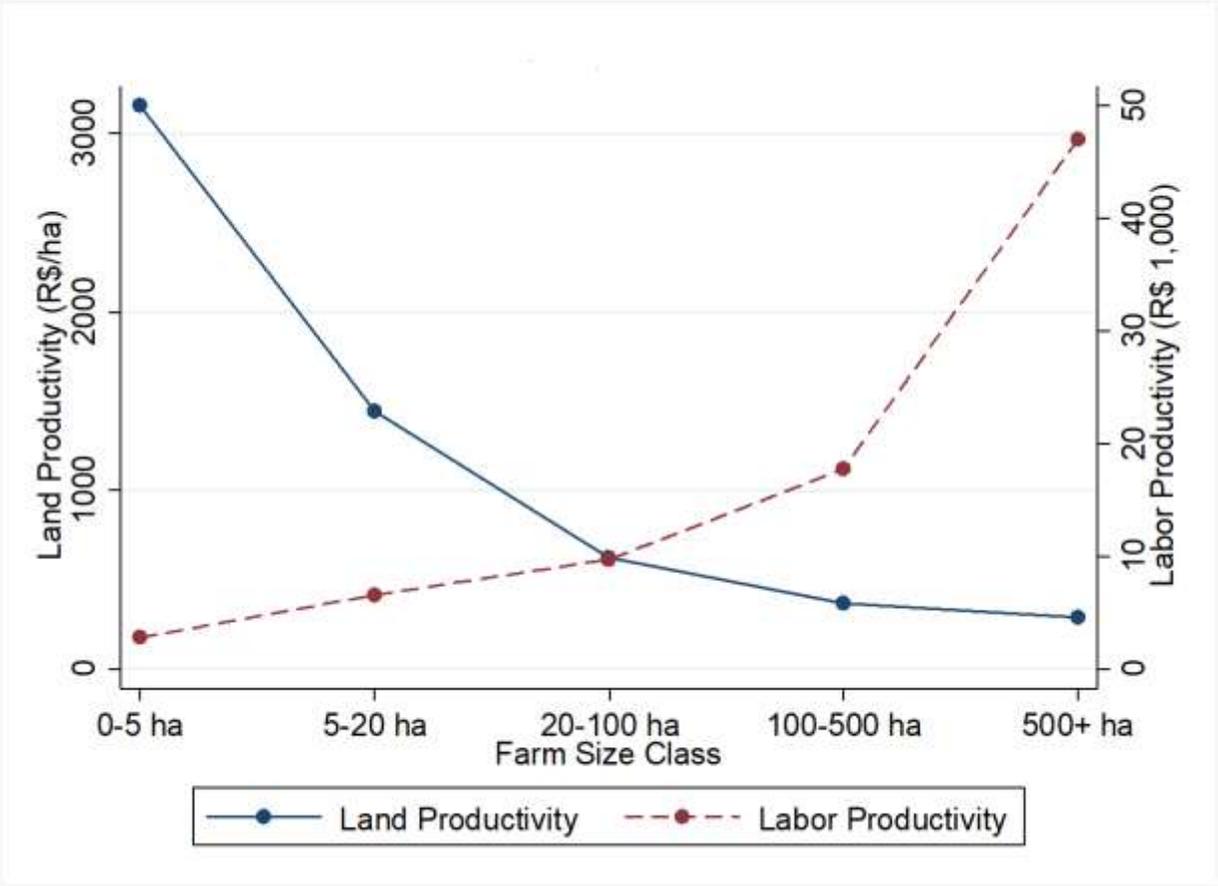
Figure 3: Land Productivity in Brazil by Region

(a) North
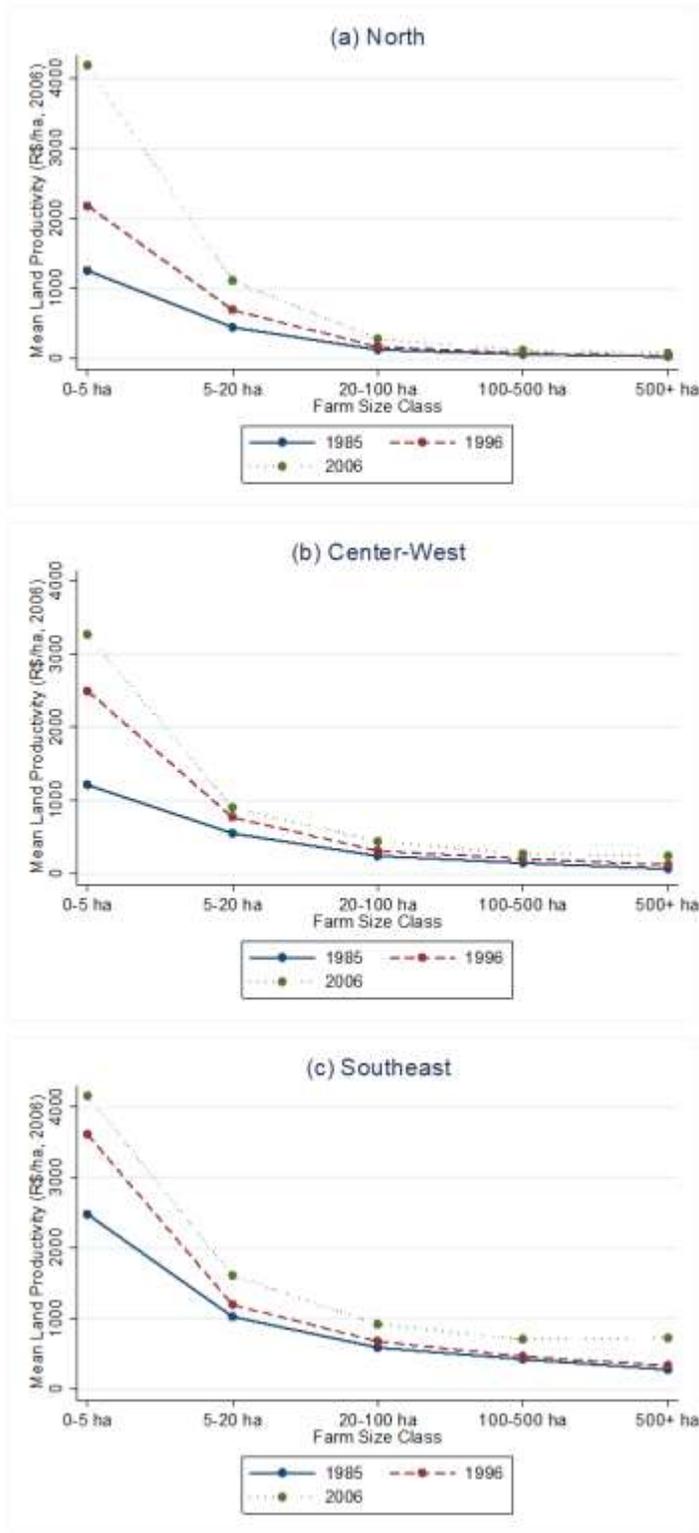
(b) Center-West

(c) Southeast

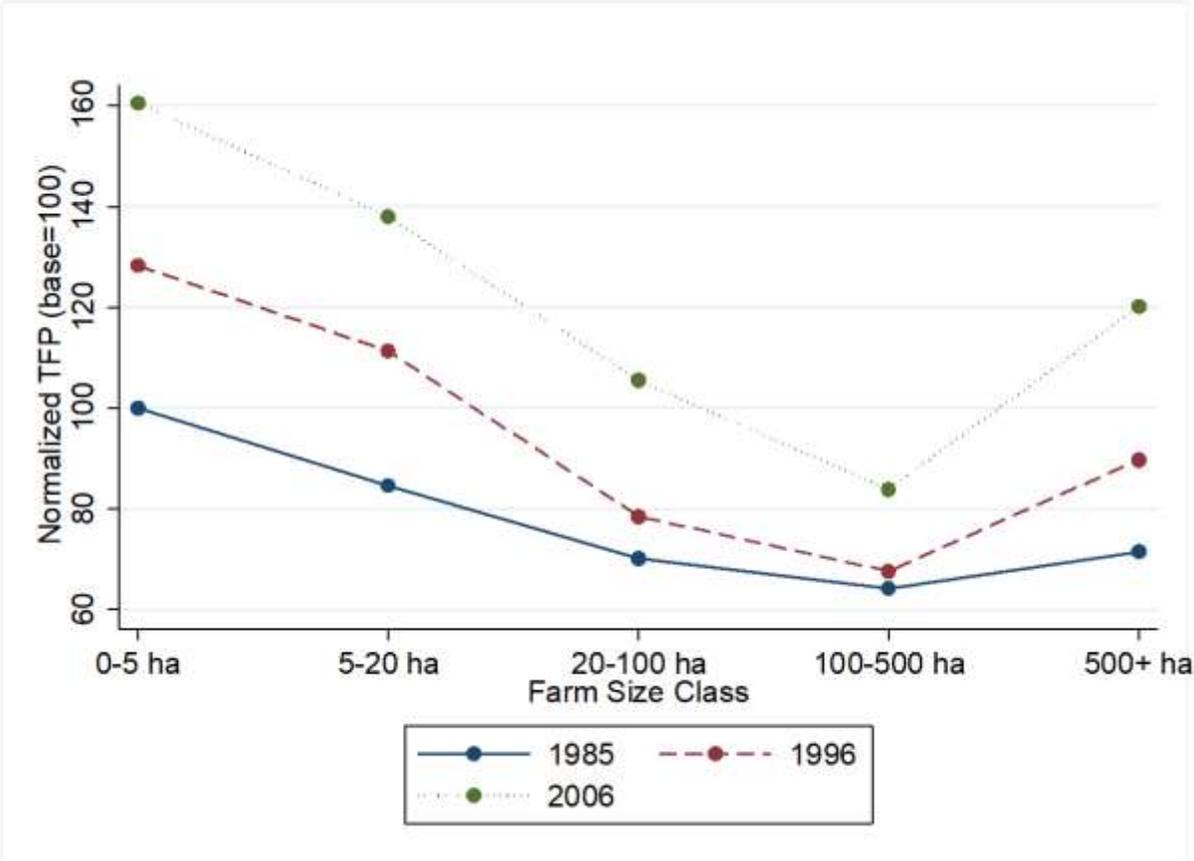**Figure 4: Total Factor Productivity in Brazil's North**

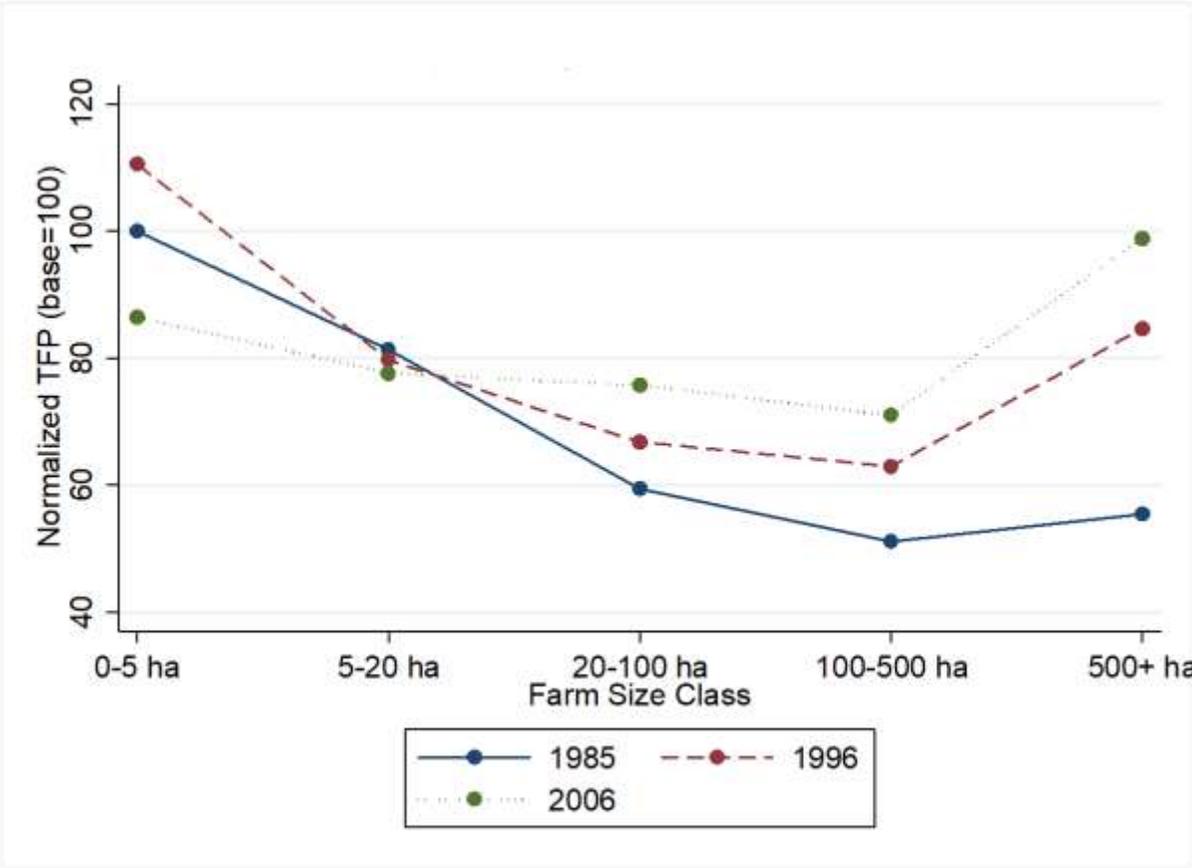**Figure 5: Total Factor Productivity in Brazil's Center-West**
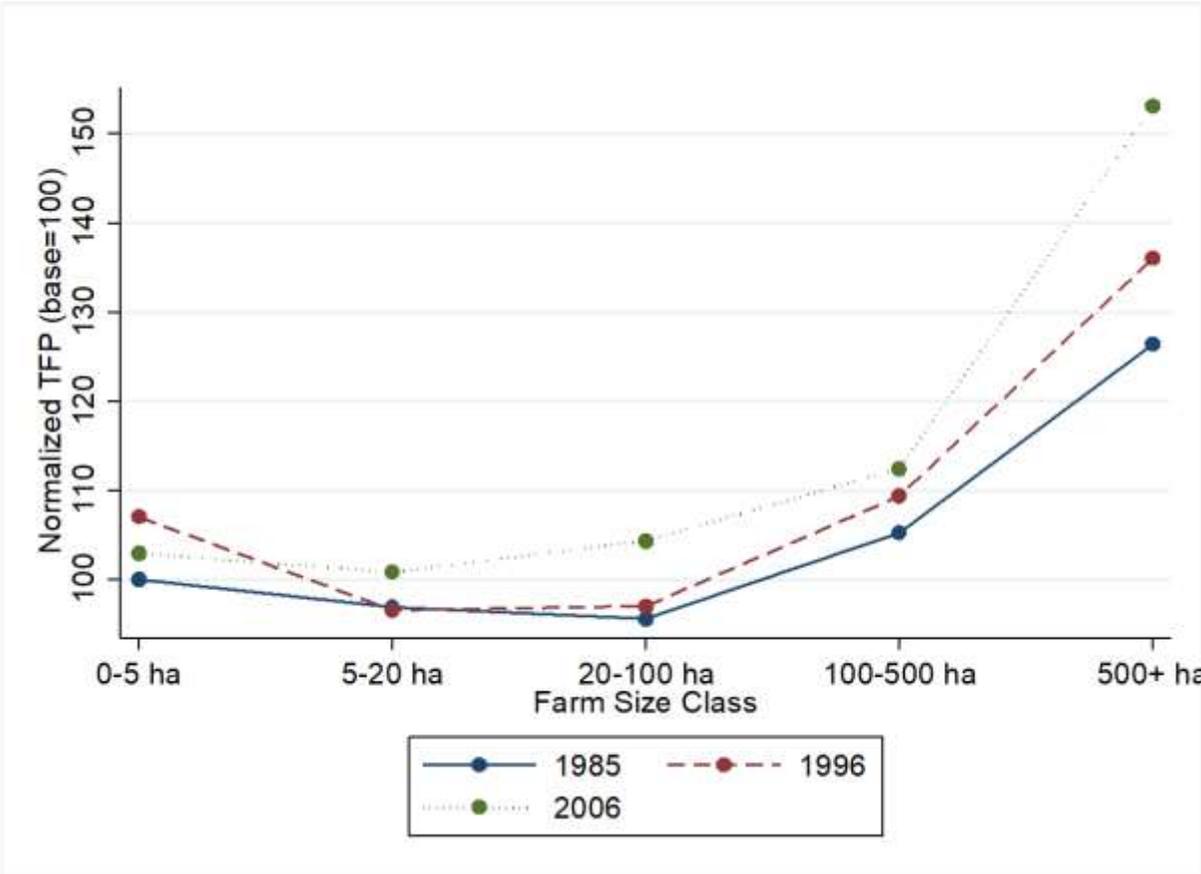
Figure 6: Total Factor Productivity in Brazil's Southeast

Table 1. Percentage Difference in TFP Relative to 0-5 ha Farms

	North			Center-West			Southeast		
	1985	1996	2006	1985	1996	2006	1985	1996	2006
5-20 ha	-15.39*	-13.22	-14.03	-18.63***	-27.90***	-10.20	-3.09	-9.76***	-2.06
	(0.086)	(0.131)	(0.293)	(0.004)	(0.003)	(0.440)	(0.383)	(0.009)	(0.623)
20-100 ha	-29.77**	-38.80***	-34.27**	-40.54***	-39.58***	-12.37	-4.41	-9.32	1.39
	(0.035)	(0.008)	(0.048)	(0.001)	(0.007)	(0.323)	(0.536)	(0.184)	(0.846)
100-500 ha	-35.87*	-47.27**	-47.70**	-48.88***	-43.02**	-17.86	5.26	2.19	9.21
	(0.053)	(0.013)	(0.038)	(0.005)	(0.034)	(0.405)	(0.641)	(0.850)	(0.442)
500 + ha	-28.61	-30.06	-25.13	-44.55*	-23.42	14.38	26.36	27.05	48.68**
	(0.382)	(0.383)	(0.547)	(0.094)	(0.471)	(0.695)	(0.150)	(0.147)	(0.018)
N	1,888	1,888	1,888	3,038	3,038	3,038	17,742	17,742	17,742

Base farm size bin, 0-5 ha. P-values from significance tests are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix A

Proof of expression (17).

$$\frac{\partial TFP}{\partial A} = \left(\frac{\partial q}{\partial A}\right)\left(\frac{1}{\tau(A)}\right) - \frac{\left(\frac{\partial \tau(A)}{\partial A}\right)q}{\tau(A)^2}$$

$$\frac{\partial TFP}{\partial A} = \left(\frac{\partial q}{\partial A}\right)\left(\frac{1}{\tau(A)}\right)\left[1 - \frac{\left(\frac{\partial \tau(A)}{\partial A}\right)q}{\tau(A)\left(\frac{\partial q}{\partial A}\right)}\right]$$

$$\frac{\partial TFP}{\partial A} = \left(\frac{\partial q}{\partial A}\right)\left(\frac{1}{\tau(A)}\right)\left[1 - \frac{\left(\frac{\partial \tau(A)}{\partial A}\right)\left(\frac{1}{\tau(A)}\right)}{\left(\frac{\partial q}{\partial A}\right)\left(\frac{1}{q}\right)}\right] = \left(\frac{\partial q}{\partial A}\right)\left(\frac{1}{\tau(A)}\right)\left[1 - \frac{\left(\frac{\partial \tau(A)}{\partial A}\right)\left(\frac{A}{\tau(A)}\right)}{\left(\frac{\partial q}{\partial A}\right)\left(\frac{A}{q}\right)}\right]$$

$$\frac{\partial TFP}{\partial A} = \left(\frac{\partial q}{\partial A}\right)\left(\frac{1}{\tau(A)}\right)\left[\frac{\varepsilon_{q,A}}{\varepsilon_{q,A}} - \frac{\varepsilon_{\tau(A),A}}{\varepsilon_{q,A}}\right]$$

$$\frac{\partial q}{\partial A} = \left(\frac{\partial TFP}{\partial A}\right)\tau(A)\left[\frac{\varepsilon_{q,A}}{\varepsilon_{q,A} - \varepsilon_{\tau(A),A}}\right]$$

Appendix B

Table A1: Data Cleaning and Sample Size, by Region and Farm Size Class

	Farm Size Class (ha)	N	Less Farm Size Outliers	Less Tails of Land Productivity Distribution	Less Land Productivity Outliers	Percent Dropped from Cleaning
North	0-5	315	315	307	307	1.6%
	5-20	420	420	410	403	4.0%
	20-100	459	459	449	437	4.8%
	100-500	443	443	433	433	2.3%
	500 +	323	315	307	305	3.2%
Center-West	0-5	537	537	525	520	3.2%
	5-20	619	619	605	605	2.3%
	20-100	681	681	667	667	1.2%
	100-500	672	672	658	658	1.9%
	500 +	613	595	583	581	2.4%
Southeast	0-5	3,850	3,850	3,772	3,772	1.2%
	5-20	3,991	3,991	3,911	3,911	1.6%
	20-100	4,024	4,024	3,942	3,942	2.0%
	100-500	3,896	3,896	3,818	3,818	0.6%
	500 +	2,235	2,215	2,169	2,169	0.9%
Brazil		47,365	47,281	*46,307*	46,261	2.3%

Table A2: Descriptive Statistics, 2006

	Farm Size Class (ha)	Output/ha	Capital/ha	Family Labor/ha	Purchased Inputs/ha	Share of Farms	Share of Area	Share of Output
North	0-5	4,185.02	965.03	1.75	862.37	19.4%	0.3%	6.9%
	5-20	1,110.46	333.92	0.24	214.43	17.4%	1.6%	12.0%
	20-100	282.24	129.06	0.05	73.54	43.0%	17.4%	32.2%
	100-500	120.65	88.47	0.01	58.18	16.9%	26.4%	20.9%
	500 +	78.98	54.90	0.01	72.96	3.3%	54.3%	28.0%
Center-West	0-5	3,265.49	2,628.76	0.72	1,971.49	8.5%	0.1%	0.8%
	5-20	902.03	851.41	0.16	506.95	20.2%	0.7%	2.4%
	20-100	444.53	378.38	0.04	231.87	37.8%	5.0%	8.3%
	100-500	276.48	279.76	0.01	210.40	21.2%	13.8%	14.3%
	500 +	247.08	127.28	0.01	246.53	12.3%	80.4%	74.2%
Southeast	0-5	4,152.85	3,903.90	0.85	1,892.53	28.8%	1.1%	5.5%
	5-20	1,611.97	1,797.21	0.18	1,061.63	32.1%	6.2%	11.7%
	20-100	923.89	906.85	0.04	554.65	28.5%	21.8%	23.5%
	100-500	711.38	555.44	0.01	556.69	9.1%	32.9%	27.2%
	500 +	726.50	276.13	0.01	715.38	1.5%	37.9%	32.1%

Table A3. Estimated Technology Coefficients

	North	Center-West	Southeast
Capital per ha	0.176***	0.175***	0.129***
	(0.039)	(0.039)	(0.015)
Family Labor per ha	0.303***	0.149***	0.190***
	(0.066)	(0.045)	(0.027)
Intermediate Inputs per ha	0.304***	0.426***	0.503***
	(0.039)	(0.040)	(0.021)
Constant	4.093***	3.627***	3.285***
	(0.244)	(0.294)	(0.104)
AMC FE	Yes	Yes	Yes
Farm Size Dummies	Yes	Yes	Yes
R-squared	0.95	0.90	0.91
N	1,888	3,038	17,742

Standard errors in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Dependent variable is logged output; all independent variables are logged; all variables normalized by farm size.