Nonlinear Relations Between Agricultural Productivity and Farm Size in India

by Anupama G.V., Thomas Falk, and Daniel Gregg

NONLINEAR RELATIONS BETWEEN AGRICULTURAL PRODUCTIVITY AND FARM SIZE IN INDIA

Anupama GV[1], Thomas Falk[2], and Daniel Gregg[3]

[1] Innovation Systems for the Drylands Program, International Crops Research Institute for Semi-Arid Tropics (ICRISAT), Patancheru 502324, Telangana, India. Corresponding author's contact: g.anupama@cgiar.org, +91 7036883739

[2] Innovation Systems for the Drylands Program, International Crops Research Institute for Semi-Arid Tropics (ICRISAT), Patancheru 502324, Telangana, India. Contact: t.falk@cgiar.org, +91 733 11 32469

[3] University of New England, Armidale, Australia. Contact: daniel.gregg@intersectioninnovations.com,

Paper prepared for presentation at the
31st International Conference of Agricultural Economists
20-25 August 2021
New Delhi, India

Introduction

Land access can be instrumental for increasing farmers' income (Jayne et al., 2010; Eastwood et al., 2010). Land is viewed not only as a production factor but also as a tool to gain access to credits (Deininger et al., 2018). It further affects inequality (Oyvat, 2016). Through limiting the aggregation of land, governments can reduce the risk of creating a small class of landholders that obtain rents from a much larger class of landless rural poor. Policies that address consolidation of fragmented agricultural land, or which place restrictions on the total area of landholding, are acting on a long-standing empirical puzzle amongst economists: the relationship between agricultural productivity and farm size (RAPFS) (Binswanger et al., 1995; Mazumdar, 1963; Rao, 1966; Rao, 1967; Sen, 1964; Eastwood et al., 2010; Ali & Deininger, 2015).

The concept of scale economies (Mill 1884) suggests that capital constraints and the scarcity of land would create a situation wherein the RAPFS was positive – i.e. that increasing the area of land operated under a single entity would generate improved productivity (Binswanger et al., 1995; Hazell, 2005; Eastwood et al., 2010). On the other hand, an inverse RAPFS implies that both productive efficiency and distributive justice could be achieved by restricting consolidation of land holdings. Indeed, the persistence of an inverse RAPFS in some areas has been used to argue for policies that limit consolidation of landholdings (Lipton, 2009), approaches to distributive justice that have been implemented in India for some time. Whilst the existence of an inverse RAPFS would be a fortunate outcome, it is unlikely to persist in contexts where there is substantial mechanization of work and improved functioning of markets: both characteristics of recently modernizing agricultural supply chains in India.

Many empirical analyses have indicated evidence of a persistent inverse RAPFS (e.g. Sen, 1964; Binswanger et al., 1995; Heltberg, 1998; Banerjee, 2000; Lamb, 2003; Eastwood et al., 2010;

Barrett et al., 2010; Ali & Deininger, 2015; Desiere & Jolliffe, 2018) including in the context of India (e.g. Rao, 1967; Manjunatha et al, 2013; Deininger et al., 2018). A range of explanations for the inverse RAPFS have been considered including omitted variable bias especially related to land quality (Carter, 1984; Binswanger et al., 1995; Eastwood et al., 2010; Barrett et al., 2010; Ali & Deininger, 2015), measurement errors (Ali & Deininger, 2015), market imperfections (Desiere & Jolliffe, 2018), labour market inefficiencies (Carter, 1984; Feder, 1985; Byiringiro & Reardon, 1996; Assunção & Braido, 2007), strategic and/or systematic over- or under-reporting dependent on the land size (Barrett et al., 2018; Desiere & Jolliffe, 2018; Carletto et al., 2013), and misspecification biases associated with the use of linear-in-parameters statistical methods (e.g. Assunção & Braido, 2007; Barrett et al., 2010; Ali & Deininger, 2015). Despite these efforts, the inverse RAPFS has been shown to be surprisingly persistent (Deininger et al. 2018).

In India, the inverse RAPFS has almost become an empirical regularity with studies from the 1960's and 1970's (Sen, 1964; Bardhan, 1973; Srinivasan, 1972) initially establishing an inverse relationship and with more recent support indicating it has continued at least until 2008 (Deininger et al. 2018). The Indian case has high relevance both for policy and for development outcomes. For example, India's latest Five Year Plan emphasises the joint role of land consolidation policies and restrictions on total landholdings in seeking to improve the productivity of land whilst ensuring distributive justice (GOI, 2015; Ghatak & Roy, 2007; Manjunatha et al, 2013). High income inequality is strongly associated with land access in India (Chakravorty et al., 2016; Deininger et al, 2017; Laha, 2017) driving emergent policy concerns with many Indian states having made efforts to transfer ownership rights to tenants (GOI, 2015; Ghatak & Roy, 2007). As a result of tenancy reform policies, by the end of 2010, 12.586 million tenants had received secure land titles covering 67,638 km^2 (ICAR, 2017). The reforms have however created strong barriers to leasing of agricultural land as owners become concerned about losing land through government-mandated title transfers to their tenants (GoI, 2016). As a result, the land rental market has become highly informal with poor tenure security for tenants. As this is likely to affect investment decisions, we expect that on leased plots the RAPFS has a smaller slope.

In the last two decades, there are no studies we are aware of on Indian agriculture or development in the that have employed panel-data involving an annual frequency for analysis combined with a large sample of households. Indeed, the most recent study (Deininger et al. 2018) involves only three periods across time – 1982, 1999, and 2008 – only one of which is in the last 20 years, and utilize a representation that will potentially be impacted by omitted variable bias due to the lack of inclusion of production substitutes. This indicates that the tenancy reforms reviewed above are likely being developed in an environment with substantial limits to contemporary information on the RAPFS (Deininger et al. 2018).

In this paper we present an analysis of in which we utilised uniquely detailed data that provide a modern insight into the RAPFS in India that enables improved management of measurement errors and apply methods that improve upon existing approaches that simultaneously deal with misspecification biases and omitted variables bias. Specifically, we used the Village Dynamics in South Asia (VDSA) panel dataset covering the years 2009 to 2015 for 1,129 households in 30 villages and 9 states of India to consider the RAPFS. The data used here involves unique depth and accuracy through collation of data using regular household visits (every three weeks) to report on multiple agricultural production variables minimising recall and measurement errors. The continuous data collection over five years better allows to control for short term weather and market dynamics. In addition, most of the studies on the RAPFS use household data to explain agricultural output (e.g. Binswanger et al, 1995; Gautam & Ahmed, 2018; Desiere & Jolliffe, 2018). The VDSA data, in contrast, permits conducting both output and profit analysis on the plot and household level. In this study we estimated the RAPFS using updated approaches to nonparametric methods, the partially-linear model (PLM) approach, that allows for a nonparametric representation of the RAPFS whilst controlling for other factors that affect productivity and profitability (i.e. for omitted variable bias). Our paper contributes to the literature in three main ways.

Firstly, we present a uniquely detailed examination of the RAPFS in the Indian context. Indian agriculture has undergone substantial changes since the 1980s. Deininger et al. (2018) find evidence for better functioning labor markets due to technological advances, rising wages and increased non-agricultural labor demand. This is likely to have an impact on the RAPFS which indicates the need for more contemporary information to support policy making. Foster & Rosenzweig (2011) are the first to express that the RAPFS is changing its shape in today's Indian agricultural sector. This finding was confirmed more recently by Deininger et al. (2018). Both studies use plot level data in 242 villages across 17 states collected in 1982, 1999 (2,424 households) and 2008 (8,659 households) as part of the Rural Economic Development Survey by India's National Council for the Applied Economic Research. Whilst these studies both indicate the potential for a shift in the RAPFS to a positive relationship, they are limited by the lack of more recent data and by the discrete nature of observations across time in their panel.

Secondly, we deepen the discussion on tenure reform policies by assessing the RAPFS under different forms of tenure. The objective in this setting was to identify whether there is a difference in the shape of the RAPFS between leasehold and owned plots. In particular, the presence of qualitative differences in the RAPFS would have potential implications for the speed and approach to rollout tenure reforms across India.

Thirdly, the analysis employs the PLM approach that extend the univariate nonparametric approach utilised in more recent times to model the RAPFS (e.g. Assunção & Braido, 2007; Barrett et al. ,2010; Ali & Deininger 2015). Whilst the use of univariate nonparametric methods may be helpful to consider complex nonlinearities they do not account for the relationships of other time-variant variables in the modelling of the RAPFS, thus trading off flexibility in the modelling the RAPFS for omitted variable bias. The PLM approach is a flexible additive model that combines linear parametric methods with a nonparametric component. The PLM approach allows for high-dimensionality in the linear component and high-flexibility in the nonparametric component. Thus the PLM approach generalises univariate nonparametric approaches to greatly improve accounting for omitted variable biases. Detecting nonlinearities related to farm size has policy relevance as land consolidation and ceiling policies target specific farm segments.

Data

The VDSA panel dataset was generated over a period of 40 years from 1975 to 2015 but with discrete periods of data collection. In the most recent period (2009-2014), the period used for this analysis, data were collected for a larger number of households and with vastly increased survey efforts focused on detailed data collection covering production information, GPS-measured plots, and 3-weekly household visits to record input and output data for each plot owned/leased by participants. The resultant data set covers the period 2009 and 2015 with 1,129 households participating from 30 villages in 9 states of India. Study sites were selected using a stepwise purposive sampling strategy in order to cover the agro-ecological diversity of the region. Within sites, households were grouped into land holding quartiles with 25% of households in each village randomly selected from each land holding quartile.

Data on household endowments were recorded once per year. Data on cultivation including all inputs and outputs were collected on the subplot level once every three weeks. Subplots refer to the separate cropping systems that may be used on any given plot. Given the diversity of agricultural practices in the Indian context, subplot level disaggregation provides for a more detailed view of activities that occur on a given plot across a given year. The short periodicity of data collection makes the data more accurate.

Two levels of aggregation were considered for this analysis: plot-level and household-level. For plot level analysis subplot level data were aggregated to the plot level whilst for the household-level analysis subplot level data were aggregated to the household level. The time period for aggregation in all cases was one agricultural year (1st June to 31st May). Consequently, the data has a panel structure with six years of observations (2009/10 to 2014/15) for 4,640 plots owned or managed by 1,129 households[1]. All monetary values were converted into 2009 prices using the wholesale price index as a conversion factor. We use the average exchange rate of 2009 (51 INR = 1 USD) to calculate values in 2009 USD values.

The majority of plots in this sample had an area of less than 1 hectare with an average of 0.53 ha and with half of all recorded plots being smaller than 0.4 ha. The average total operated area is 1.37 hectare with the majority of households controlling less than 1 hectare. The average profit is 419 USD and half of the households generate a profit below 277 USD. More than 90 percent of all households would live below the poverty line if agriculture was their only income (Figure 1).

[1] The panel is not balanced as sample household migrated out of the sites and were replaced by new households using the same sampling procedure.

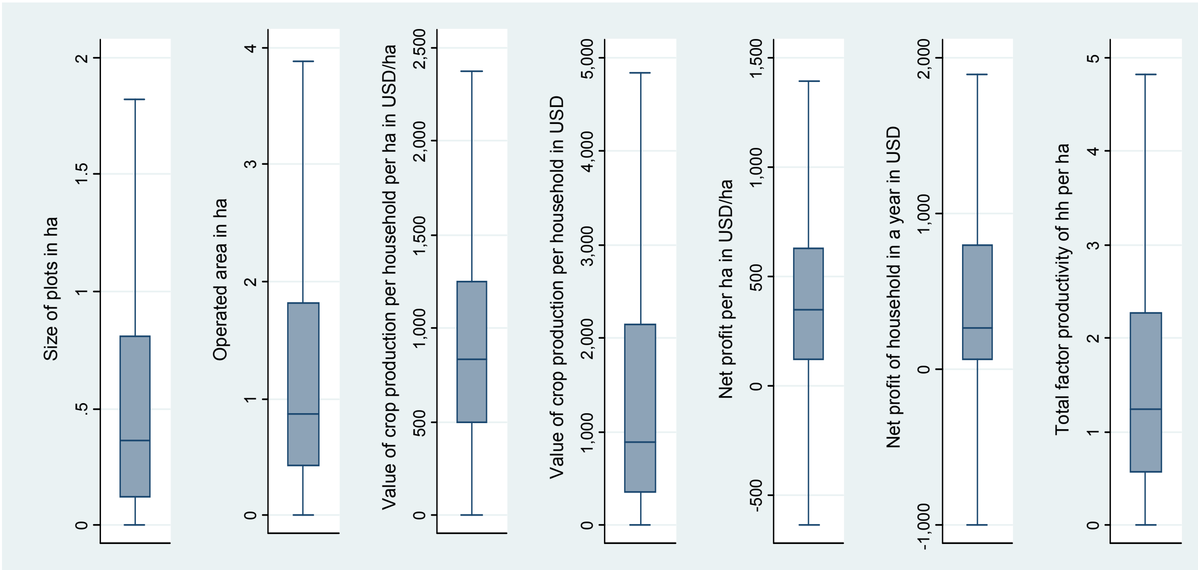

Figure 1. Boxplots of key farm related indicators (outside values not shown)

Model Specification

Our models build on well-established economic theory in this field which has been summarized by Barrett et al. (2010). They take into account household fixed effects and plot characteristics which were identified as key factors when studying the RAPSF. Equation 1 represents the theory underlying our analysis:

$$y_{ij} = \gamma \cdot m(A_{ij}) + \beta' x_{ij} + \emptyset\, q_{ij} + \alpha' z_i + \nu s_i + \lambda_i + \delta t + \omega_{ij} \quad (1)$$

The function, $m(A_{ij})$, provides a linkage between two estimation approaches considered here: (1) a parametric/linear model, and; (2) a semiparametric/partially linear model (Robinson 1988). In the case of the linear model $\gamma \cdot m(A_{ij}) = \gamma \cdot A_{ij}$. In the case of the partially linear model $m(A_{ij})$ was estimated using the Robinson (1988) double residual method (see Verardi & Debarsy 2012).

The dependent variable y_{ij} is the agricultural productivity per hectare of household *i* on plot *j*. A_{ij} is the key variable of interest representing the size of the plot which was cultivated in any season of that year. The vector x_{ij} contains variables that are associated with production at the *jth* plot (e.g. labour, material, water inputs). $\emptyset\, q_{ij}$ is a vector of observed plot characteristics including soil depth as an indicator for land quality and information on cropping systems; s_i is the vector of observed socioeconomic controls; z_i includes household-level time-varying factors whilst λ_i represents household fixed effects. δt provides for time fixed effects in the specification. Errors for the additive linear function are assumed symmetric IID. The household-level function structurally corresponds to the plot-level function but with plot-level factors aggregated to the household level with associated dropping of *j* plot subscripts.

Three indicators are used to measure agricultural productivity y_{ij}:

a) Total value of crop output obtained per hectare over all seasons in a year (crop value);
b) Net profit per hectare over all seasons in a year (profit);
c) Total factor productivity (TFP) per ha.

The crop value of a plot is the total amount of the crop main product and by-product in all the seasons of the specific agricultural year multiplied by the prevailing harvest prices at the time of harvest in the respective village).

The plot level net profit was calculated by deducting the costs of all inputs applied to the plot during the year from the crop value of the plot for the year. Input and labour costs were calculated on the basis of reported quantities and location- and time-specific prices. Inputs include electricity, fuel, seeds, organic and inorganic fertilizers, pesticides, fungicides, growth regulators, micro nutrients, weedicides, bullocks, tractor, thresher, harvester, and other machinery costs. Labour includes hired and family male, female, and child labour. In the case of family labour and use of owned machinery, shadow prices were included in the cost calculation. For the specific type of machinery and labour applied, the time of use was multiplied by location- and time-specific prices reflecting how much the household would have had to pay if they had hired the machinery or the labour. The first principal component has a strong positive load on all these indicators. The household level net profit was calculated using the same logic, deducting the costs of all inputs applied by the household during the year from the crop value of the household for the year.

Total factor productivity per ha at the plot level is computed by dividing the crop value of the plot for the year by all inputs applied to the plot during the year. The same input variables as for the net profit calculation have been used. The household level total factor productivity is the crop value of the household for the year divided by all inputs applied by the household during the year.

All metric variables were transformed using the natural logarithm or inverse hyperbolic sine transformation in case of variables containing also negative values (Burbidge *et al*. 1988). We excluded extreme observations which fall outside three times the interquartile range of any of the variables in the dependent variable series (yij) or in the area variable (Aij). This criterion was fulfilled for 832 or 5.68 percent of the plot level observations.

The parametric models were estimated using Ordinary Least Squares. The semipar function (Robinson, 1988) in STATA 14 was used to estimate Robinson's semiparametric regression models.

Results

We will present the results of the parametric and semiparametric models by providing for each productivity indicator a figure showing the parametric and semiparametric fit. In addition, we present tables with the key model information relevant for understanding the RAPFS. The first sub-section of the results is dedicated to the plot level and the second to the household level analysis. A third sub-section provides results on land tenure related effects.

Plot level analyses

Figures 2a and 2b indicate a statistically significant positive RAPFS at the plot level both for the crop value as well as the profit (also Table 1). The nonparametric component of the semi-parametric model indicates a relatively linear relation for the crop value and the total factor productivity. In the case of the net profit, we see that productivity increases most strongly as very small plots gain in size.

a) CROP VALUE (*Table 1, Models 1.1 & 1.2; parametric fit significant at 0.1% level*)
b) NET PROFIT (*Table 1, Models 1.3 & 1.4; parametric fit significant at 0.1% level*)
c) TFP *(Table 1, Model 1.5 & 1.6; parametric fit significant at 0.1% level)*

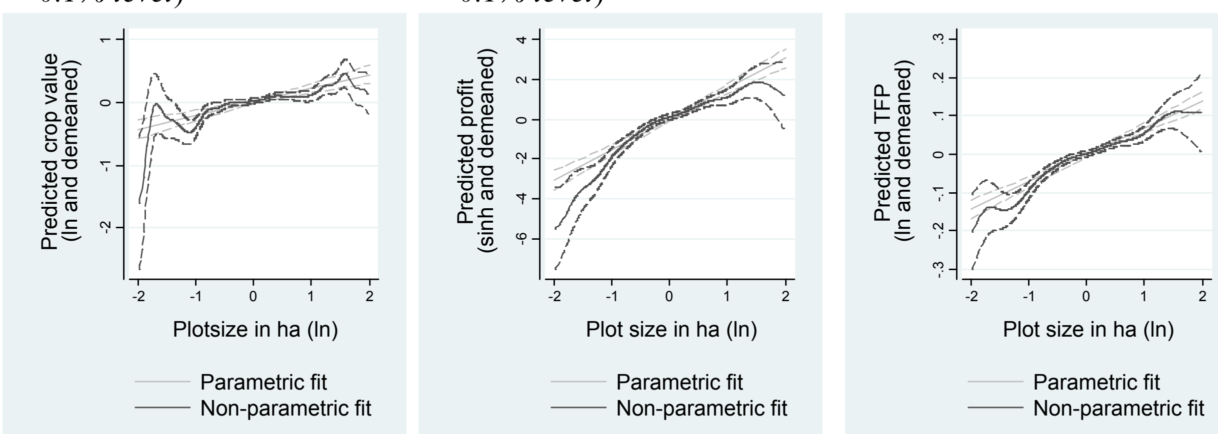

Figure 2. Parametric and semi-parametric functions explaining agricultural productivity indicators on the plot level using parametric and Robinson's semiparametric regression estimator. The dotted lines show the confidence intervals.

Referring to the deep literature on the RAPFS one could suspect that we may have found a negative relation if we had not controlled for market failures and plot characteristics. We therefore replicate the analysis on the basis of the simplest specification presented in Barrett et al. (2010) which disregards household fixed effects and plot characteristics. Even these models show a consistent positive RAPFS.

Table 1. Parametric and Robinson's semi-parametric models explaining the relation between plot size and plot level productivity indicators with household fixed effects. Coefficients with standard errors in parentheses, ** $p < 0.01$, *** $p < 0.001$.

	ln CROP VALUE		asinh NET PROFIT		Ln TFP	
	Semi-parametric	Para-metric	Semi-parametric	Para-metric	Semi-parametric	Para-metric
	Model 1.1	Model 1.2	Model 1.3	Model	Model 1.5	Model 1.6
Ln plot size in ha	*Figure 2a*	0.217*** (0.038)	*Figure 2b*	1.526*** (0.117)	*Figure 2c*	0.071*** (0.006)
Ln operated area in ha	-0.143* (0.068)	-0.121 (0.068)	-0.239 (0.309)	-0.126 (0.309)	-0.011 (0.015)	-0.007 (0.015)
Observed production inputs (x_{ij})	yes	yes	no	no	no	no
Observed plot characteristics (q_{ij})	yes	yes	yes	yes	yes	yes
Observed socioeconomic controls (s_i)	yes	yes	yes	yes	yes	yes
Year dummies(t)	yes	yes	yes	yes	yes	yes
Constant		-0.367*** (0.100)		-1.156** (0.429)		0.027 (0.021))
Adj. R^2	0.196	0.197	0.021	0.039	0.034	0.050
Observations	12508	12508	12508	12508	12490	12490
Log likelihood	-22254	-22305	-41099	-41116	-3602	-3614

Notes: Likelihood ratio tests of significance indicate that village dummies do not significantly improve the model fit**.**

Household level analyses

The household level analyses confirm the same trend, though the results are less clear. Specifically, the parametric model does not show a significant relation between crop value and the households operated land area (Figure 3a, Table 2 Model 2.2) whilst the models explaining the profit show a significant positive relation. The non-parametric model indicates that the net profit increase is strongest for smaller farms and stagnates for bigger ones (Figure 3b, Table 2 Model 2.3 &2.4).

a) CROP VALUE (*Table 2, Models 2.1 & 2.2; parametric fit not significant*)

b) Net PROFIT (Table 2, Models 2.3 & 2.4; *parametric fit significant at 1% level*)

c) TFP (Table 2, Models 2.5 & 2.6; *parametric fit significant at 0.1% level*)

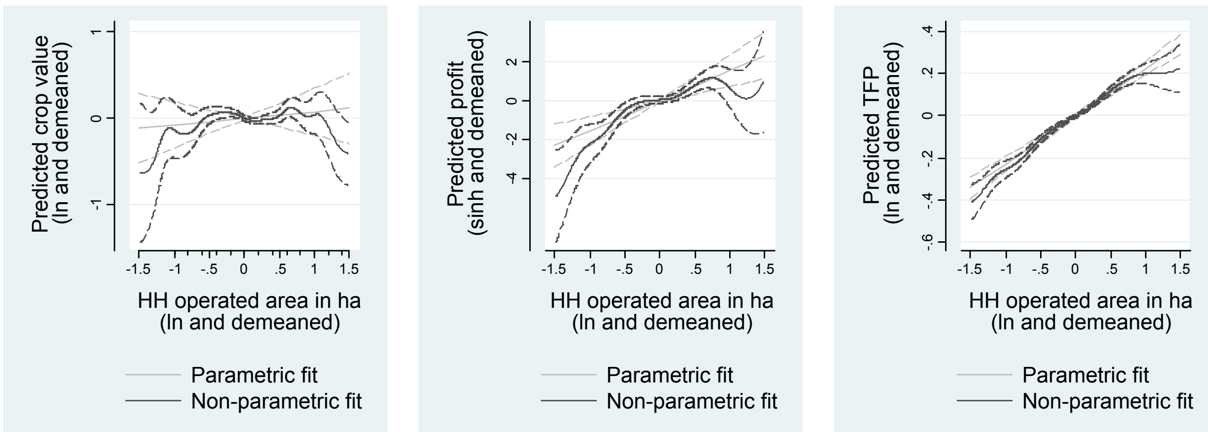

Figure 3. Parametric and semi-parametric functions explaining the relation between household operated area and household agricultural productivity indicators using parametric and Robinson's semiparametric regression estimator. The dotted lines show the confidence intervals.

The nonparametric component of the semi-parametric model shows a very linear relation between the operated area and the total factor productivity. The results confirm that the total factor productivity decreases if the operated are is split into many plots. The models disregarding household fixed effects and land quality indicators show an even stronger positive RAPFS.

Table 2. Parametric and Robinson's semi-parametric models explaining the relation between households' operated area and household level productivity indicators with household fixed effects. Coefficients with standard errors in parentheses, *** $p < 0.001$.

	ln CROP VALUE		asinh Net PROFIT		Ln TFP	
	Semi-parametric	Para-metric	Semi-parametric	Para-metric	Semi-parametric	Para-metric
	Model 2.1	Model 2.2	Model 2.3	Model 2.4	Model 2.5	Model 2.6
Ln household's operated area in ha	Figure 3a	0.097 (0.120)	Figure 3b	1.545^{***} (0.357)	Figure 3c	0.227^{***} (0.016)
Ln number of plots	-0.009 (0.021)	-0.018 (0.024)	0.018 (0.099)	-0.048 (0.097)	-0.043^{***} (0.010)	-0.045^{***} (0.010)
Observed production inputs (x_{ij})	yes	yes	no	no	no	no
Observed land characteristics (q_{ij})	yes	yes	yes	yes	yes	yes
Observed socioeconomic controls (s_i)	yes	yes	yes	yes	yes	yes
Year dummies(t)	yes	yes	yes	yes	yes	yes
Constant		0.452 (0.272)		2.654 (2.455)		0.521^{***} (0.114)
Adj. R^2	0.125	0.125	0.032	0.044	0.049	0.136
Observations	5444	5444	5272	5272	5272	5272
Log likelihood	-7400	-7453	-16238	-16259	-345	-356

Notes: Likelihood ratio tests of significance indicate that village dummies do not significantly improve the model fit.

Land tenure related Subsample Analysis

In 2014, 14 percent of the surveyed plots where leased. The analysis of subsamples for leased and owned plots show a consistent picture of positive RAPFS. There is little difference in the shape of the crop value function between leased and owned plots (Figures 4a & 4b) indicating that tenure had little relationship to the RAPFS.

(a) m(Ai) for CROP VALUE estimated for LEASED plots only *(parametric fit significant at 1% level)*

(b) m(Ai) for CROP VALUE estimated for OWNED plots only *(parametric fit significant at 0.1% level)*

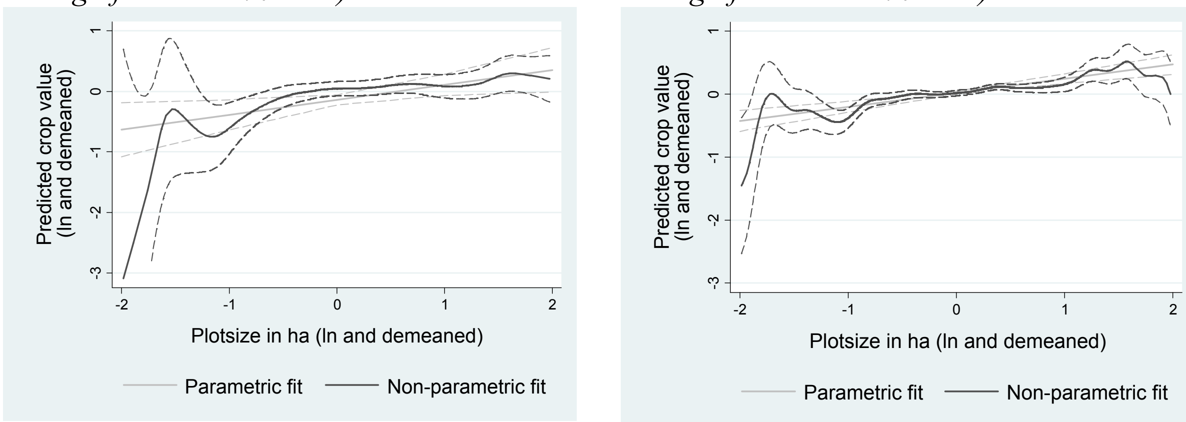

Figure 4. Comparing the RAPFS of (a) LEASED and (b) OWNED plots based on the indicator plot-level crop value per hectare. Parametric and Robinson's semi-parametric functions with household fixed effects are presented.

The net profit models indicate a significant and positive RAPFS in the case of leased plots that is undifferentiated across plot sizes. In the case of the owned plots, however, smaller plots have a higher marginal improvement to increases in plot area than larger plots (Figure 5a and Figure 5b).

(a) RAPFS for profit per hectare estimated for LEASED plots only *(parametric fit significant at 1% level)*

(b) RAPFS for profit per hectare estimated for OWNED plots only (*parametric fit significant at 0.1% level*)

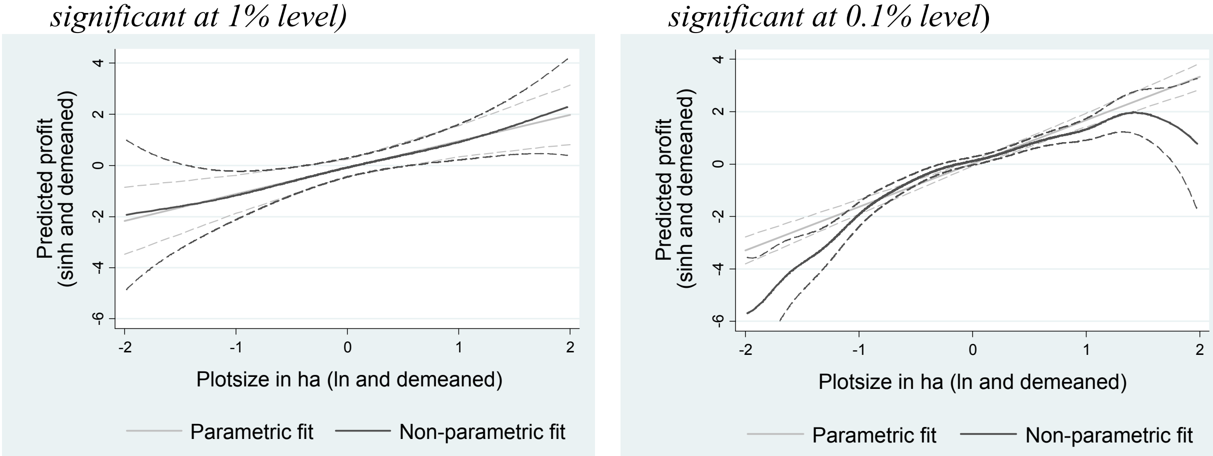

Figure 5. Comparing the RAPFS of (a) LEASED and (b) OWNED plots based on the indicator plot-level net profit per ha. Parametric and Robinson's semi-parametric functions with household fixed effects are presented.

Discussion

Our analysis provides consistent evidence for a positive RAPFS in Indian smallholder agriculture in the 2010s. No matter whether we control for household fixed effects and plot characteristics, whether we use crop value or profit as productivity indicator, or whether we conduct plot or household level analysis, the farm size coefficients are positive and highly significant in almost all cases, the exception being household level analyses in which the estimated RAPFS is insignificant for a large portion of the range. The contrast of our findings of a consistently positive RAPFS against a relatively large, but also largely out dated, body of evidence indicating an inverse RAPFS indicates that policy frameworks around tenancy reform and equity objectives associated with retaining smallholdings in agricultural systems in India may need reviewing. In particular, a positive RAPFS indicates that alternative approaches to equity objectives that are more direct may be more effective and efficient for achieving development objectives (i.e. facilitating aggregation of land with sellers moving into alternative livelihood pathways).

These results differ from many previous studies for the case of Indian agriculture (Binswanger et al., 1995; Lamb, 2003; Assunção & Braido 2007). Whilst data collection and time-period differences may be a difference, with our study using an approach that limits measurement errors, there is also the possibility that the situation for Indian agriculture has simply changed in more recent times. This conclusion would be in line with the analysis of Foster and Rosenzweig (2011) and Deininger et al., (2018). The latter argue that the shape of the RAPFS has changed due to changes in wage levels, newly available technologies and non-agricultural labor demand. Between 2006 and 2014, wage rates in India have increased by more than three times (ILO 2016b) with the result that capital increasingly substitutes for labour. New information technologies make it easier to supervise hired labour. As the result, it is very likely that the productivity disadvantages of hired labour which negatively affects the productivity of larger farms decreased over the last decades (Deininger et al., 2018). Our models indicate a still large efficiency difference between family and hired labour. It can be questioned whether markets will ever be able to balance the motivational effect of working on your own farm.

In line with the most recent evidence (Deininger et al., 2018) from 2008, the results show that investing in material inputs has a considerably more positive effect on the productivity than investing in hired labour on a per-rupee basis. This supports the hypothesis of factor substitution. If this was the case, land ownership should become increasingly important as it potentially facilitates credit access (Deininger et al., 2018). Theoretically, owners of large plots, compared to tenants of large plots, would have more financial capital and stronger incentives to invest in their land with greater potential to achieve economies of scale (Eastwood et al., 2010). Our results do not indicate any substantive and/or significant differences in the RAPFS for owned versus leased plots. This may partially be due to the analysis involving mostly smaller plots here where even the 'larger' plots in this sample may not be sufficient in size to achieve true economies of scale.

Whilst there has been a substantial focus on the potential for nonparametric approaches to assist in resolving the 'paradox' associated with a persistently inverse RAPFS, our results show no such promise. In all cases the RAPFS estimated here is positive, and in most cases significantly so. The remaining key differences between this analysis and previous analyses are: (1) omitted variable bias for estimation of univariate nonparametric functions; (2) aggregation level (we include plot-level analysis); (3) the time period of analysis – ours being up to 6 years later than the most recent analysis in India and having 6 years of data (compared to only 1 in the next most recent comparable study), and; (4) data quality. Whilst we are unable to test (4), the differences in data quality, our study provides insights into the remaining three potential causes of our findings of a positive RAPFS.

Our results do not show any support for the potential for omitted variable bias reasoning with univariate nonparametric methods and partial linear approaches both indicate a strongly positive and significant RAPFS at the plot level.

The second case, aggregate analysis at the household level indicates substantial potential differences from plot-level analysis however. Specifically, estimation of the RAPFS at the household level (aggregating all plots owned by households) indicated that the RAPFS was not significantly different from zero across a large range of farm sizes for either revenue or profit measures. Given the large number of studies that rely on a household-level analysis this provides a likely candidate for the RAPFS being 'consistently' negative over the 40 years of analysis in India, and in other regions.

The third case is also a potential cause of differences with Deininger et al (2018) finding that the inverse RAPFS had become substantially less negative over time. Given the more recent dynamism of agriculture and food supply chains in India, including indications in this study and that labour markets have become more efficient over time, our findings support those of Deininger et al (2018).

Conclusion

Our study tested for the direction of the relationship between agricultural productivity and farm size (RAPFS). In contrast to a large number of other studies, including studies based also in India, we find strong evidence for a positive RAPFS. We suggest these differences may be associated with genuine changes in the RAPFS that have shifted the relationship from a weak negative one

(Deininger et al. 2018) to a strongly positive one in recent times. A range of approaches confirmed the robustness of this approach to omitted variables bias and functional form restrictions including the use of alternative productivity measures, inclusion of land quality measures, and the application of semi-parametric methods. The positive relationship is observed on the plot and the household level as well as for three different productivity indicators.

The presence of a positive RAPFS implies trade-offs between food production and reducing inequalities (Eastwood et al., 2010; Harris & Orr, 2014). This demands more sensitive policy choices due to the potential for land consolidation policies to generate improved productivity but potentially also increases in inequality (Deininger et al., 2018). The results also indicate that productivity related to operated land area is undifferentiated between freehold and leasehold land indicating that land consolidation need not be based on permanent transfers of land (Thapa & Niroula, 2008).

However, there appears to be a relatively low ceiling for improving household level productivity. Our results show that profit increases associated with increasing household operated land area are effectively zero for moderately large plots. Thus, preventing large land agglomerations with the aim to improve wealth distribution as intended by land ceiling policies may cause only moderate productivity loss (NABARD, 2018).

Reference

Ali, D. A. & Deininger, K. (2015). Is there a farm size–productivity relationship in African agriculture? Evidence from Rwanda. *Land Economics, 91*(2), 317-343. doi.org/10.3368/le.91.2.317

Assunção, J. J., & Braido, L. H. (2007). Testing household-specific explanations for the inverse productivity relationship. *American journal of agricultural economics*, *89*(4), 980-990. doi.org/10.1111/j.1467-8276.2007.01032

Banerjee, A. V. (2000, December). Prospects and strategies for land reform. In *Annual World Bank Conference on Development Economics 1999* (pp. 253-84). World Bank, Washington D.C. dx.doi.org/10.2139/ssrn.183711

Bardhan, P. K. (1973). Size, productivity, and returns to scale: An analysis of farm-level data in Indian agriculture. *Journal of Political Economy*, *81*(6), 1370-1386. doi.org/10.1086/260132

Barrett, C. B., Bellemare, M. F., & Hou, J. Y. (2010). Reconsidering conventional explanations of the inverse productivity–size relationship. *World Development*, *38*(1), 88-97. doi.org/10.1016/j.worlddev.2009.06.002

Barrett, C., Abay, K., Abate, G., & Bernard, T. (2018). Correlated non-classical measurement errors, second best policy inference and the inverse size-productivity relationship in agriculture. Paper presented at the 30th International Conference of Agricultural Economists, July 28th – August 2nd 2018, Vancouver/Canada. doi.org/10.1016/j.jdeveco.2019.03.008

Binswanger, H. P., & Jodha, N. S. (1978). Manual of instructions for economic investigators in ICRISAT's Village Level Studies. Village Level Studies Series Economic Program v. 2. ICRISAT, Patancheru, India.

Binswanger, H. P., Deininger, K., & Feder, G. (1995). Power, distortions, revolt and reform in agricultural land relations. *Handbook of development economics*, *3*, part B, 2659-2772. doi.org/10.1016/S1573-4471(95)30019-8

Burbidge, J. B., Magee, L., & Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, *83*(401), 123-127. doi.org/10.1080/01621459.1988.10478575

Byiringiro, F., & Reardon, T. (1996). Farm productivity in Rwanda: effects of farm size, erosion, and soil conservation investments. *Agricultural economics*, *15*(2), 127-136. doi.org/10.1016/S0169-5150(96)01201-7

Carletto, C., Savastano, S., & Zezza, A. (2013). Fact or artefact: the impact of measurement errors on the farm size-productivity relationship. *Journal of Development Economics, 103*, 254-261. doi.org/10.1016/j.jdeveco.2013.03.004

Carter, M. R. (1984). Identification of the inverse relationship between farm size and productivity: an empirical analysis of peasant agricultural production. *Oxford Economic Papers*, *36*(1), 131-145.

Chakravorty, S., Chandrasekhar, S., & Naraparaju, K. (2016). Income generation and inequality in India's agricultural sector: The Consequences of land fragmentation (No. 2016-028). Indira Gandhi Institute of Development Research, Mumbai, India.

Deininger, K., Monchuk, D., Nagarajan, H. K., & Singh, S. K. (2017). Does land fragmentation increase the cost of cultivation? Evidence from India. *The Journal of Development Studies*, *53*(1), 82-98. doi.org/10.1080/00220388.2016.1166210

Deininger, K., Jin, S., Liu, Y., & Singh, S. K. (2018). Can labor market imperfections explain changes in the inverse farm size–productivity relationship? Longitudinal evidence from rural India. *Land Economics*, *94*(2), 239-258. doi.org/10.3368/le.94.2.239

Desiere, S., & Jolliffe, D. (2018). Land productivity and plot size: Is measurement error driving the inverse relationship? *Journal of Development Economics, 130*, 84-98. doi.org/10.1596/1813-9450-8134

Eastwood, R., Lipton, M., & Newell, A. (2010). Farm size. *Handbook of agricultural economics*, *4*, 3323-3397. doi.org/10.1016/S1574-0072(09)04065-1

FAO. (1996). AEZ map of the developing world. FAO-UN - Land and Water Division (CBL) FAO, Rome, Italy.

Feder, G. (1985). The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints. *Journal of development economics*, *18*(2-3), 297-313. doi.org/10.1016/0304-3878(85)90059-8

Foster, A. D., & Rosenzweig, M. R. (2011). Are Indian farms too small? Mechanization, agency costs, and farm efficiency. *Unpublished Manuscript, Brown University and Yale University*.

Gautam, M., & Ahmed, M. (2018). Too small to be beautiful?: the farm size and productivity relationship in Bangladesh. *Food Policy (2018)*. doi.org/10.1016/j.foodpol.2018.03.013

Ghatak, M., & Roy, S. (2007). Land reform and agricultural productivity in India: a review of the evidence. *Oxford Review of Economic Policy*, *23*(2), 251-269. doi.org/10.1093/oxrep/grm017

Government of India (GOI). (2015). Five Year Plans. Planning Commission. Yojana Bhavan, Sansad Marg, New Delhi - 110001.

Government of India (GOI). (2016). Report of the Expert Committee on Land Leasing. National Institution for Transforming India. New Delhi/India.

Harris, D., & Orr, A. (2014). Is rainfed agriculture really a pathway from poverty? *Agricultural Systems*, *123*, 84-96. doi.org/10.1016/j.agsy.2013.09.005

Hazell, P. B. (2005). Is there a future for small farms? *Agricultural Economics*, *32*(s1), 93-101. doi.org/10.1111/j.0169-5150.2004.00016.x

Heltberg, R. (1998). Rural market imperfections and the farm size—productivity relationship: Evidence from Pakistan. *World Development*, *26*(10), 1807-1826. doi.org/10.1016/S0305-750X(98)00084-9

ICAR (2017). Handbook of Agriculture. Indian Council of Agricultural Research (ICAR), Govt. of India, New Delhi Edition: 6th edition 2013, reprint Nov.2017 ISBN: 9788171640966

International Labour Office. (2016a). Global Wage Report 2016/17: Wage inequality in the workplace. ILO.

International Labour Organization, J. (2016b). India Labour Market Update. ILO Country Office for India. New Delhi, India

Jayne, T. S., Mather, D., & Mghenyi, E. (2010). Principal challenges confronting smallholder agriculture in sub-Saharan Africa. *World development*, *38*(10), 1384-1398. doi.org/10.1016/j.worlddev.2010.06.002

Laha, A. (2017). Tenancy and inequality in land ownership in India: Evidence from household surveys. *IASSI-Quarterly*, *36*(1), 61-74. Print ISSN : 0970-9061. Online ISSN : 0974-018X.

Lamb, R. L. (2003). Inverse productivity: Land quality, labor markets, and measurement error. *Journal of Development Economics*, *71*(1), 71-95. doi.org/10.1016/S0304-3878(02)00134-7

Lipton, M. (2009). Land reform in developing countries: Property rights and property wrongs. Routledge. ISBN: 1134863144, 9781134863143

Manjunatha, A. V., Anik, A. R., Speelman, S., & Nuppenau, E. A. (2013). Impact of land fragmentation, farm size, land ownership and crop diversity on profit and efficiency of irrigated farms in India. *Land Use Policy*, *31*, 397-405. doi.org/10.1016/j.landusepol.2012.08.005

Mazumdar, D. (1963). "On the economics of relative efficiency of small farmers." Econ. Weekly, Jul. 1963, pp. 1259-63. - . "Size of Farm and Productivity: A Problem of Indian Peasant Agriculture." *Economica 32*(1965): 161-73

Mill, J. S. (1884). Principles of political economy. D. Appleton.

NABARD (2018). All India rural financial inclusion survey 2016-17. National Bank for Agriculture and Rural Development. Mumbai/India.

Oyvat, C. (2016). Agrarian structures, urbanization, and inequality. *World Development*, *83*, 207-230. doi.org/10.1016/j.worlddev.2016.01.019

Rao, A. P. (1967). Size of holding and productivity. *Economic and Political Weekly*, Vol. 2, No.44, Nov. 1967, pp. 1989-1991.

Rao, C. H. H. (1966). Alternative explanations of the inverse relationship between farm size and output per acre in India. *Indian Economic Review*. Vol. 1, No.2, (1966):1-12.

Rao, Y.M., Chand, K.R., Kiresur, V.R., Deb, U.K.D., & Bantilan, M.C.S. (2015). Documentation of Village Dynamics Studies (VDSA) in South Asia (2009-11), Patancheru 502 324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics. 145 pp.

Robinson, P. M. (1988). Root-N-consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, Vol. 56, No.4 (Jul., 1988), pp. 931-954. Published by The Econometric Society. doi.org:/10.2307/1912705.

Sen, A.K. (1964). Size of holdings and productivity. *Economic Weekly*, Feb. 1964, Annual Number 16. pp. 323-6.

Srinivasan, T. N. (1972). Farm size and productivity implications of choice under uncertainty. *Sankhyā: The Indian Journal of Statistics, Series B*, *(1960-2002)* Vol. 34, No. 4 (Dec., 1972), pp. 409-420. Published by: Indian Statistical Institute.

Thapa, G. B., & Niroula, G. S. (2008). Alternative options of land consolidation in the mountains of Nepal: An analysis based on stakeholders' opinions. *Land use policy*, *25*(3), 338-350. doi.org/10.1016/j.landusepol.2007.09.002

Thorat, A., Vanneman, R., Desai, S. and Dubey, A. (2017). Escaping and falling into poverty in India today. *World development*, 93: 413-426. doi.org/10.1016/j.worlddev.2017.01.004

Verardi, V. & Debarsy, N. 2012. Robinson’s square root of N consistent semiparametric regression estimator in Stata. *The Stata Journal*, 12(4): 726-735. doi.org/10.1177/1536867X1201200411

Walker, T. S., & Ryan, J. G. (1990). *Village and household economics in India's semi-arid tropics*. Johns Hopkins University Press. Baltimore. 394 pp. ISBN 0-8018-3886-X