Land Distribution and International Agricultural Productivity

Dietrich Vollrath

The unequal distribution of agricultural land is often cited as a source of inefficiency in agriculture. Previous cross-country studies of agricultural productivity differences, though, have not considered land inequality. This article addresses this issue by using cross-country data on inequality in operational holdings of agricultural land from Deininger and Squire (1998). In an estimation of an agricultural production function, the Gini coefficient for land holdings is found to have a significant negative relationship with productivity. This is consistent with the existence of heterogeneity in productivity by farm size within countries. A one standard deviation drop in the Gini coefficient implies an increase in productivity of 8.5%.

Key words: agricultural productivity, agricultural production function, land inequality.

A wide range of studies have considered the cross-country productivity of agriculture (Hayami and Ruttan 1970; Kawagoe, Hayami, and Ruttan 1985; Lau and Yotopoulos 1988; Fulginiti and Perrin 1993; Frisvold and Ingram 1995; Craig, Pardey, and Roseboom 1997; Mundlak 2000).[1] These works have attempted to quantify the role that elements such as capital, land quality, infrastructure, and research and development have in determining the variation in agricultural output between countries. None of these studies, though, have addressed the distribution of agricultural land and its place in agricultural productivity.

This omission is somewhat surprising given the number of theoretical connections drawn between land distribution and agricultural productivity as well as the relevance of the cross-country evidence for discussions of land reform and land policy. This article examines the relationship of land distribution and agricultural productivity across countries by utilizing data on the distribution of operational farm size within countries calculated by Deininger and Squire (1998). As an introduction, consider figure 1 which plots output per hectare against land inequality as measured by the Gini coefficient. There is a significant negative relationship, showing that inequality in operational holding size within a country is associated with low productivity.

The robustness of the relationship in figure 1 is addressed by including the Gini coefficient in the estimation of an aggregate agricultural production function. The results show that the negative relationship between land inequality and productivity persists even when controlling for aggregate input use, land quality, human capital, agricultural research effort, and other country-specific institutional factors. The estimates of the production elasticities for agricultural inputs are in line with previous work in this area.

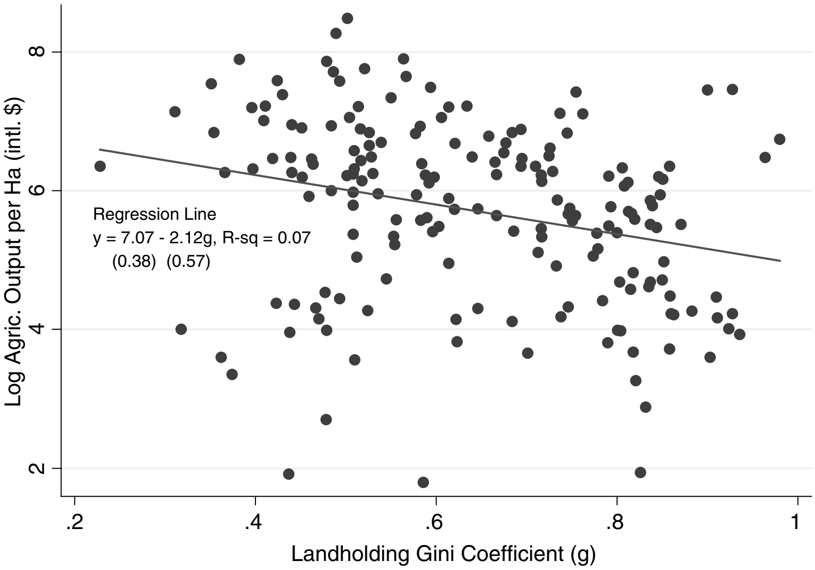
The negative relationship between land distribution and productivity is consistent with the productivity advantages of farms operated primarily with family labor, something documented by several lines of research. A series of works (Johnston and Kilby 1975; Johnston and Clark 1982; Tomich, Kilby, and Johnston 1995) examine the difference between *unimodal* (or equitable) and *bimodal* (or unequal) agrarian structures. They stress that for most countries the unimodal structure is more productive because it equalizes the marginal product of labor across farms. Labor misallocations arise in bimodal structures because labor supervision costs and policy distortions combine to make capital relatively cheap for large farms.

The advantages of the unimodal structure are based in part on the observed inverse farm-size productivity relationship (IFSP).

Dietrich Vollrath is assistant professor, Department of Economics, University of Houston. The author thanks Aimee Chin, Keith Wiebe, the editor (Christopher B. Barrett), and two anonymous referees for their valuable comments. All remaining errors are the author's.

[1] This article follows a common convention in the literature and uses the term *productivity* to refer to the partial productivity of land (i.e., output per hectare) as opposed to total factor productivity.

Amer. J. Agr. Econ. 89(1) (February 2007): 202–216



Source: Deininger and Squire (1998) and FAO (2000)

Figure 1. Agricultural output and land distribution

Evidence of the IFSP relationship has been found in a variety of settings (Berry and Cline 1979). Decreasing returns to scale in agriculture is generally rejected as an explanation (Bardhan 1973; Berry and Cline 1979; Carter 1984) and the evidence suggests that variation in the shadow price of labor drives the IFSP relationship (Carter 1984). Several authors (Bardhan 1973; Feder 1985; Eswaran and Kotwal 1986; Frisvold 1994; Heltberg 1998) focus on supervision costs as the source of this variation. Feder (1985) further shows that multiple market failures are necessary to make shadow prices vary by farm size. In these cases, a more equitable distribution of operated holdings would improve aggregate efficiency.

Alternative theories of the IFSP relationship, however, do not lead to this same prediction. Bhalla (1988) and Bhalla and Roy (1988) find that including measures of soil quality in their specifications removes the IFSP result. Benjamin (1995) uses an instrumental variable approach to show that omitted variables (presumed to be related to land quality) explain the IFSP relationship. As well as controlling for land quality, Lamb (2003) also suggests that the evidence points to measurement error of farm size as being a source of the IFSP relationship. A further possibility unrelated to land quality is that an economy consisting of households facing uncertainty regarding prices will display an IFSP relationship even without any market imperfections or measurement issues (Barrett 1996). In these cases, a change in the distribution of operated holdings would not necessarily lead to higher productivity.

There are several individual country studies that explore the possible connection of land distribution and productivity. Besley and Burgess (2000) finds that land reforms in India were associated with lower poverty and higher agricultural wages, but that this occurred through changes in production relationships rather than through changes in land distribution. They do find that land reforms had their greatest effect in those Indian states with the greatest initial land inequality. Jeon and Kim (2000) document significant productivity gains from the land reforms undertaken in Korea in the 1950s which limited the amount of land any individual could own. Examining the historical nature of land relationships in India, Banerjee and Iyer (2005) find that those Indian states with higher initial land inequality had lower productivity even after land reform legislation took effect.

If there are such advantages to more equitable distributions of land, it begs the question of why land rentals and sales market have not allocated land more efficiently. The failure of land markets are well documented. Land sales markets may fail to bring about an efficient distribution of operational holdings for several reasons, such as covariate risk, imperfect credit markets, and policies that raise the price of land above the capitalized value of the agricultural output (Binswanger, Deininger, and Feder 1995). Within rental markets, sharecropping is often used in place of more efficient fixed-rent contracts due to incentive issues (Otsuka, Chuma, and Hayami 1992). Legal restrictions and the threat of land reform may also act to limit rental market activity (de Janvry and Sadoulet 1989; Horowitz 1993; Diaz 2000).

Regardless of the reason, the results here suggest that the failure of land markets to allocate land efficiently has serious aggregate consequences. The point estimates in this article show that a one standard deviation fall in the Gini coefficient is associated with an increase in productivity of 8.5%. Looking at this another way, the difference in median Gini coefficients for Latin America, 0.81, and the members of the Organisation for Economic Co-operation and Development (OECD), 0.56, is associated with a difference in output per hectare of approximately 13%, holding constant aggregate input use, land quality, human capital, and institutional quality.

The article proceeds as follows. The next section describes the land distribution data of Deininger and Squire in more detail. The following section presents the estimation of the aggregate production function and the last section concludes.

Measuring Land Distribution

The distribution of land holdings is measured using data from Deininger and Squire (1998) (DS hereafter). The authors computed Gini coefficients for the size distribution of land holdings within countries using data obtained from the decennial agricultural censuses of the Food and Agriculture Organization of the U.N. (FAO). These data provide totals of both the number of holdings and total area of holdings, broken down by size of holdings. From this a Lorenz curve is estimated that can be used to calculate the Gini coefficient. The size classes used are standard across countries and years, with some exceptions, so that the Gini is comparable across countries.

The Gini coefficient very distinctly measures the distribution of *operational* holdings. For the purposes of their census taking, the FAO defined an agricultural holding as, "... an economic unit of agricultural production under single management comprising all livestock kept and all land used wholly or partly for agricultural production purposes, without regard to title, legal form, or size." (FAO 1997, p. 13) Thus, the Gini coefficient does not capture the distribution of land *ownership* within a country. For the purposes of this article, though, it is precisely the distribution of operational holdings that is relevant because we are interested in efficiency, not equity.

Table 1. Median Operational Land Holding Gini Coefficient, by Region

Region	Median Gini
Sub-Saharan Africa	0.49
East Asia	0.51
OECD	0.56
South Asia	0.59
Eastern Europe	0.62
Mideast and North Africa	0.66
Latin America	0.81

Note: Author's calculations using data from Deininger and Squire (1998).

DS provide 286 observations across 117 countries ranging in time from 1939 to 1994.[2] The mean Gini coefficient is 0.64 with a standard deviation of 0.16. The lowest observed Gini coefficient is 0.23 (Sweden in 1971) and the highest is 0.98 (Hungary in 1980). Table 1 shows the median Gini coefficient by region for the whole sample. Perhaps not unsurprisingly, Latin American countries tend to have higher inequality in land distribution than other regions. The low median value for sub-Saharan African countries suggests that low land inequality does not necessarily lead to high agricultural productivity.

One limit of the Gini coefficient is that it cannot distinguish between differences in the scale of agricultural holdings across countries. Consider a country consisting of a handful of enourmous plantations that are all of a similar size. This country would have a very low Gini coefficient and it would not be possible to distinguish it in the data from a country consisting of thousands of small family farms. To address this, an additional control for average farm size is constructed from summary data in the 1990 World Census of Agriculture (FAO 1997).[3]

Average farm size ranges from a high of 3,601 hectares in Australia in 1990 to a low of

[2] DS originally report 261 observations across 103 countries. The latest version of their data set contains additional observations. They do not specify which observations have been added since the publication of their paper.

[3] The FAO reports total number of holdings as well as total area of holdings and so average holding size is simply calculated. It must be noted, though, that the total area of holdings does not necessarily correspond to the total agricultural area reported by the FAO. This is because the agricultural censuses from which the

Table 2. Median Holding Size, by Region

Region	Median Holding Size (Ha)
East Asia	2.07
Sub-Saharan Africa	2.18
South Asia	2.32
Mideast and North Africa	6.05
Eastern Europe	8.69
Latin America	17.70
OECD	20.83

Note: Author's calculations using data from the Report on the 1990 World Census of Agriculture, FAO, 1997.

0.73 hectares in Turkey in 1980. Table 2 shows the median holding size by region. There exists a ten-fold difference in holding size between the OECD and the regions with the smallest holdings (East Asia, Sub-Saharan Africa, and South Asia). The OECD group includes Australia, Canada, New Zealand, and the United States, all of which have very high average holding size. Even excluding those countries, though, the median holding size in the OECD falls only to 17.74 hectares. The Latin American median is pulled down by relatively small holding sizes in the Carribean nations. For countries in South America itself, the median holding size increases to 68.70 hectares.

A plot of the Gini coefficient and the log of average holding size is shown in figure 2. There is a small but statistically significant positve relationship between the two measures. Both measures will be included in each specification to capture both aspects of land distribution.

Land Distribution and Agricultural Productivity

To begin with it will be useful to discuss what the cross-country evidence is able to tell us. Figure 1 shows a negative relationship between land inequality and productivity, but this correlation does not identify why land distribution matters. To be more explicit, consider the following simple decomposition of agricultural productivity. A portion of all farms, $1-\lambda$, are small farms, while the remaining portion, λ, are large farms. The term λ is thus a crude proxy for the Gini coefficient of land inequality. The average land area of small farms is given by θ_s while the average land area of the large farms is θ_l, and $\theta_l > \theta_s$.

holding area is derived are in some cases samplings from the full population of holdings.

Overall output per hectare in the economy, y, is simply a weighted average of the output per hectare of each farm type

$$(1) \qquad y = A[(1-\lambda)\theta_s f_s(\mathbf{x}_s) + \lambda\theta_l f_l(\mathbf{x}_l)]$$

where A is total factor productivity (TFP). The terms $f_s(\mathbf{x}_s)$ and $f_l(\mathbf{x}_l)$ are the per hectare production functions applicable to small and large farms, respectively, and $\mathbf{x}_i$, $i = s, l$ is the vector of per hectare inputs used by each type of farm. If there is no difference in production between the types of farms, then $f_s(\mathbf{x}_s) = f_l(\mathbf{x}_l)$ and the expression for output per hectare in (1) reduces to

$$(2) \qquad y = A[(1-\lambda)\theta_s + \lambda\theta_l]f(\mathbf{x})$$

where $f(\cdot)$ is the general production function common to both kinds of farms and $\mathbf{x}$ is the vector of aggregate input use per hectare. The term in brackets in equation (2) is simply average farm size in a country. Looking across countries, if both TFP (A) and average farm size are controlled for, then we would not observe any relationship between the distribution of land (λ) and output per hectare (y).

On the other hand, if we do observe a relationship between distribution and output, then this implies differences between the types of farms in output per farm. More succinctly, any observed relationship of λ and y, again controlling for TFP and average farm size, implies that $\theta_s f_s(\mathbf{x}_s) \neq \theta_l f_l(\mathbf{x}_l)$ in equation (1). In addition, if the relationship is *negative*, then this implies further that $\theta_s f_s(\mathbf{x}_s) > \theta_l f_l(\mathbf{x}_l)$. Combined with the original assumption that $\theta_l > \theta_s$, then it follows that $f_s(\mathbf{x}_s) > f_l(\mathbf{x}_l)$. Thus, an observed negative relationship between land inequality and output per hectare implies a productivity advantage for small farms. This is consistent with the market-failures version of the IFSP relationship and, to the extent that small farms are equivalent to family operated farms, it is consistent with the literature on the unimodal agrarian structure.[4]

The approach of this article is therefore to estimate an agricultural production function, controlling for both average farm size as well

[4] Note that the land quality explanation of the IFSP relationship would predict no clear relationship between y and λ. This is because both $f_s(\mathbf{x}_s)$ and $f_l(\mathbf{x}_l)$ are assumed to be related to λ as well. Any decrease in inequality (λ going down) would be associated with $f_s(\mathbf{x}_s)$ decreasing as poorer land was put to use by small farms. So, there would not necessarily be an increase in output per hectare.

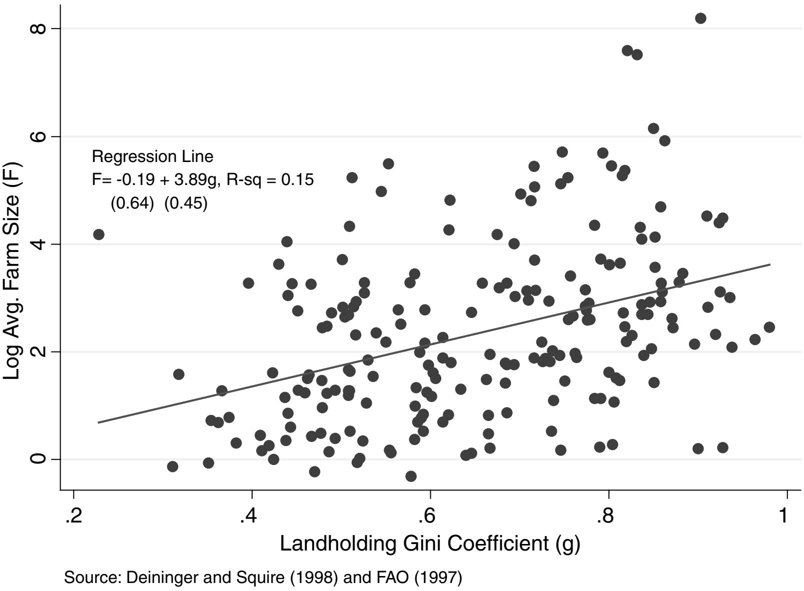


Figure 2. Land distribution and average holding size

as for elements affecting total factor productivity. These elements include land quality, human capital, and the general institutional environment. If the Gini coefficient for land holdings (of which λ is a proxy) is still significant after controlling for these things, then this offers evidence that there is heterogeneity in productivity by farm size. Note that this cannot say anything about *why* there is heterogeneity in productivity by farm size, only that it has an empirically significant impact on aggregate productivity.

In estimating the effect that the distribution of operational holdings has on productivity, a basic assumption will be that all countries share a common production function. This assumption is common to the literature on cross-country agricultural productivity. The specification used for estimation follows this literature as well and can be written in its most general form as

$$(3) \qquad \ln y_i = \beta_0 + \beta_1 g_i + \beta_2 \ln l_i + \beta_X \ln X_i + \beta_Z Z_i + \epsilon_i$$

where y_i is output per hectare, g_i is the Gini coefficient, l_i is land per holding, X_i is a vector of inputs in per hectare terms, Z_i is a vector of other country-specific control variables including land quality, and ϵ_i is a potentially heteroskedastic error term. The coefficients β_1 and β_2 capture the partial productivity effect of land inequality and average farm size, respectively. β_X and β_Z are vectors of coefficients for their respective vectors of control variables, and β_0 is a constant.

The measures of g_i and l_i have already been discussed. The measure of output is the total value of all agricultural production after deductions for feed and seed. Land is measured as the total hectares of agricultural land. Descriptions of the sources of these and all other data are included in Appendix A.

The vector X_i consists of four inputs commonly used in the estimation of agricultural production functions: livestock, tractors, fertilizer, and labor. Land, of course, is also an input to agricultural production. Excluding total land from X_i implicitly assumes the production function is constant returns to scale. Including total land in X_i would allow for the possibility of decreasing or increasing returns. The results in this article are not affected by this decision and only the constant returns to scale results are presented.

Within Z_i, there are several categories of controls. The first involves the quality of labor and is controlled for by both life expectancy and total fertility rate.[5] The second is land quality, which is controlled for by the percent of land irrigated as well as a land quality index

[5] Additional data regarding primary school enrollment rates from the World Bank's Development Indicators were examined as well. Due to the highly colinear relationship with both life expectancy and total fertility rates, it did not add any meaningful information to the regressions, but did greatly reduce the sample size.

from Peterson (1987). Additionally, the percent of total land given to permanent pasture is included to control for an effect observed by Berry and Cline (1979) across countries. They found that large farms put less of their total area under cultivation and the remainder is either unused or given over to pasture. They also find that the relative underutilization of land becomes more severe the more unequally distributed is the land.

Engerman and Sokoloff (1997) and Sokoloff and Engerman (2000) propose that land inequality has influenced the general quality of institutions within countries, focusing in particular on those countries in the Western Hemisphere. According to their work, concentrated land distributions allowed a landholding elite to control the political system and they limited both political participation and education reform in order to provide a relatively cheap agricultural labor force. Binswanger and Deininger (1997) also discuss the avenues by which large landholders can skew the institutional environment to their favor. To control for the interaction of land distribution and institutions, an institutional quality index is included as a control. In addition, variables regarding the origins of the legal system are included as this may influence the functioning of land markets within countries as well as playing a part in initial land inequality in ownership.

A final category of controls regards agricultural research. Craig, Pardey, and Roseboom (1997) and Masters (2003) include agricultural research and development (R&D) as an input into agricultural production and show that it is positively associated with productivity. To control for the possibility that land inequality might be acting as a proxy for research effort, the level of expenditure on agricultural R&D is included as an input into the production function.

Ordinary Least Squares Specifications

The combination of the Gini coefficient and average holding size data with the data on productivity and the other control variables results in a sample of 177 observations ranging in time from 1958 to 1993. While the FAO organizes the agricultural census decennially, the actual year in which the census is conducted by a country varies greatly. This results in observations that cover twenty-nine different years and eighty separate countries. Of these, twenty-four are observed only once, while one country (France) has five observations. Table 3 presents summary statistics for all the variables.

The base estimations pool the 177 observations together and uses ordinary least squares (OLS) to estimate specifications of the form found in (3). Table 4 presents the results of these regressions. The initial specification in column (1) includes only the Gini coefficient and average holding size as controls. The Gini is negatively related with output per hectare, and is significant at the 5% level. The point estimate indicates a very strong correlation of land inequality and productivity, with a one standard deviation decrease in the Gini coefficient associated with an increase in output per hectare of 23%.

Column (2) adds controls for input use and the most obvious result is that the point

Table 3. Summary Statistics for OLS Regression Data

Variable	Observations	Mean	SD	Min	Max
Output per ha (intl. $)	177	574.33	691.44	6.00	4,847.75
Gini coefficient	177	0.64	0.16	0.23	0.98
Avg. farm size (ha)	177	77.11	340.23	0.73	3,601.68
Livestock per 1,000 ha (cow equivalents)	177	681.42	546.97	19.41	3,475.10
Fertilizer per 1,000 ha (tons)	177	64.85	86.73	<0.01	357.39
Tractors per 1,000 ha (number)	177	19.23	40.11	0.01	376.29
Labor per 1,000 ha (number)	177	369.55	527.89	0.91	3,025.73
Land quality index	177	106.38	43.88	27.00	249.00
Percent land irrigated	177	8.97	14.93	<0.01	74.26
Percent land in pasture	177	47.37	29.72	0.85	99.22
Total fertility rate (children per woman)	177	4.28	2.14	1.26	9.34
Life expectancy (years)	177	62.95	11.09	34.92	78.84
Institutions index	177	0.40	0.87	−1.75	1.74
Agric. R&D per 1,000 ha (PPP $)	121	23.71	105.28	<0.01	1,076.31

Note: See Appendix for sources. The institutions index is from Kaufman, Kraay, and Zoido -Lobaton (2002).

Table 4. OLS Regression Results

	Dep Variable: Log Agricultural Output per Hectare (intl. $)						
Exp Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gini coefficient - g_i	−1.63**	−0.77**	−0.70**	−0.55**	−0.71**	−0.59**	−0.73**
	(2.42)	(3.38)	(3.22)	(2.60)	(3.41)	(2.83)	(3.56)
Log avg. farm	−0.08	−0.01	−0.02	−0.02	−0.02	−0.02	0.06
size − $\ln l_i$	(1.37)	(0.20)	(0.37)	(0.65)	(0.37)	(0.56)	(0.84)
Inputs							
Log livestock per ha		0.30**	0.31**	0.32**	0.31**	0.34**	0.26**
		(4.68)	(5.06)	(5.21)	(5.20)	(5.03)	(2.99)
Log fertilizer per ha		0.20**	0.19**	0.19**	0.14**	0.14**	0.13*
		(3.81)	(3.81)	(3.96)	(2.70)	(3.06)	(1.87)
Log tractor per ha		0.19**	0.19**	0.17**	0.15**	0.13**	0.08**
		(4.43)	(4.73)	(4.63)	(4.26)	(3.86)	(2.63)
Log labor per ha		0.18**	0.12**	0.08	0.19**	0.13**	0.26**
		(2.84)	(2.12)	(1.61)	(2.96)	(2.36)	(2.28)
Land quality							
Peterson LQI			0.32**	0.26**	0.30**	0.21**	0.13
			(3.09)	(2.75)	(3.21)	(2.48)	(1.17)
% Irrigated			0.15	0.01	0.22	−0.12	0.09
			(0.56)	(0.03)	(1.00)	(0.48)	(0.35)
% Permanent pasture				−0.42**		−0.66**	−0.30
				(2.31)		(3.06)	(1.11)
Research effort							
Log agric. R&D							0.09**
expend per ha							(3.40)
Constant	6.93**	3.98**	3.47**	3.51**	3.47**	3.46**	4.29**
	(15.76)	(8.34)	(7.78)	(7.64)	(5.15)	(4.63)	(4.31)
Human capital and[a] *institutional controls*	No	No	No	No	Yes	Yes	Yes
Observations	177	177	177	177	177	177	121
R-squared	0.07	0.90	0.91	0.91	0.92	0.93	0.94

Notes: Absolute values of t-statistics are given in parentheses. Single asterisk(∗) denotes significance at 10%, double asterisk(∗∗) denotes significance at 5%. Standard errors are robust to heteroskedasticity.

[a]Specific controls for human capital are total fertility rate and life expectancy. Controls for institutions include the Kraufman, Kraay, and Zoido-Lobaton (2002) index and dummies for legal origin from La Porta et al. (1999)

estimate for the Gini coefficient is more than halved, but with increased significance. This result is generally replicated in column (3) when the land quality index from Peterson (1987) and the percent of land irrigated are controlled for as well. In column (4), the percent of land used as pasture is added as a control and this has a more significant impact on the estimates. The estimated effect of the Gini coefficient drops again, indicating a correlation of land inequality and the type of agriculture being pursued. This is in line with the discussion in Berry and Cline (1979) regarding the relative underutilization of land in highly unequal countries.

Columns (5) and (6) include controls for human capital, institutions, and legal origins. As can be seen, these do not materially affect the estimated effect of the Gini coefficient. Again, the point estimate on the Gini drops when the percent of pasture land is controlled for but remains significant. The size of the point estimate on the Gini coefficient remains practically significant as well. Column (6) indicates that a one standard deviation drop in the Gini coefficient is associated with an increase in productivity of 10%.

The final column of table 4 includes the log of agricultural research and development expenditures per hectare as a control. The limitations of these data shrink the sample to only 121 observations, but the results are generally comparable to the previous regressions. The point estimate for the Gini coefficient is relatively large considering that the percent of pasture land is controlled for as well. The coefficient on R&D expenditures is significantly positive, and there is some difference in the point estimates for the other inputs into production. In particular, the land quality index is no longer significant and has a much smaller point estimate.

Overall, table 4 shows that the relationship of operational land inequality and productivity is negative and significant, even after controlling for aggregate input use, human capital, institutions, and research effort. Assuming that this successfully controls for all the possible level effects of land inequality on productivity, the significant estimate of the Gini coefficient implies that there is heterogeneity of productivity by farm size. This is consistent with broad-based or unimodal distributions of land having efficiency advantages. This result is also consistent with the presence of an IFSP relationship within countries.

The fact that some countries do not achieve a more equitable distribution of operational holdings presumably has much to do with the functioning of land sales and rental markets within those countries. Estimates from these regressions offer a way to quantify the potential gains from land market reforms or land redistributions. For example, consider the distribution of land in Argentina in 1988. Holdings of less than 200 hectares constituted 74.5% of all holdings, but only controlled 7.7% of the land (FAO, 1997, p. 147). If reforms were undertaken that allowed this share of land to increase to 15%, the Gini coefficient in Argentina would drop from 0.85 to 0.80. Given the estimates in column (6) of table 4, this would be associated with an increase in output per hectare of 3%. This effect, keep in mind, presumes that there would be no change in the aggregate input use in Argentina, nor would the percent of land used for pasture change. So the full effect on the economy may in fact be much larger than this estimate implies.

Panel Specifications

The nature of the Gini coefficient data from DS limited the size of the available data set for the simple OLS regressions in the previous section. There may be some concern that these results are not comparable to previous estimates of agricultural production functions that utilize a much longer time series of data. To address this, the values of the Gini and average holding size are interpolated and the production function is estimated by panel techniques.

There are fifty-four countries which have multiple observations of the Gini coefficient and average holding size. For those countries, values for the Gini and average holding size are interpolated for years lying between observed years. Data on output, agricultural inputs, and the percentage irrigated and in pasture are available on a yearly basis from the FAO and are used directly. The control data that are time varying (total fertility rate, life expectancy, and agricultural R&D) are interpolated when necessary, while the other control data are time-invariant (land quality, institutions, and legal origins). Specific details on the interpolation technique can be found in the Appendix. The interpolation results in an unbalanced panel of 1,159 observations, with the number of observations per country ranging from six to thirty-two. Table 5 summarizes the variables.

The specification in (3) is modified to be

$$\ln y_{it} = \beta_0 + \beta_1 g_{it} + \beta_2 \ln l_{it} + \beta_X \ln X_{it} + \beta_W W_{it} + \beta_Z Z_i + \upsilon_i + \omega_t + \epsilon_{it} \tag{4}$$

Table 5. Summary Statistics for Panel Regression Data

Variable	Observations	Mean	SD	Min	Max
Output per ha (intl. $)	1,159	624.36	733.41	17.80	4,847.75
Gini coefficient	1,159	0.65	0.17	0.23	0.98
Avg. farm size (ha)	1,159	108.32	405.03	0.73	3,601.67
Livestock per 1,000 ha (cow equivalents)	1,159	721.29	550.68	63.25	3,475.1
Fertilizer per 1,000 ha (tons)	1,159	76.09	98.34	<0.01	481.42
Tractors per 1,000 ha (number)	1,159	21.11	38.18	0.01	376.29
Labor per 1,000 ha (number)	1,159	349.46	478.58	0.89	2,570.60
Land quality index	54	107.02	41.58	54.00	249.00
Percent land irrigated	1,159	8.61	14.11	<0.01	65.30
Percent land in pasture	1,159	46.07	28.82	0.85	94.37
Total fertility rate (children per woman)	1,159	3.93	2.02	1.26	9.34
Life expectancy (years)	1,159	64.96	9.95	34.92	78.84
Institutions index	54	0.49	0.88	−1.75	1.74
Agric. R&D per 1,000 ha (PPP $)	1,011	15.85	51.04	<0.01	1,076.31

Note: See Appendix for sources. The institutions index is from Kaufman, Kraay, and Zoido-Lobaton (2002).

where the variables are now indexed by both i (for country) and t (for time period). The control variables are divided into those that are time-varying (W_{it}) and those that are time-invariant (Z_i). υ_i is a country-specific constant, ω_t is a year-specific constant, and ϵ_{it} is an error term that will be addressed in several different ways.

The first specifications is fixed-effects. In this case, ϵ_{it} is a potentially heteroskedastic error term, and v_i is allowed to be correlated with g_{it}, l_{it}, X_{it}, and W_{it}. The fixed-effect estimation uses first differences of (4) and so both the υ_i term and the Z_i term fall out. Columns (1) and (2) of table 6 show the results of fixed-effects regressions, differing only in whether the percent of pasture land is included. In both columns, the Gini coefficient is found to be significant and with a smaller point estimate than the comparable OLS estimates. Unlike the OLS estimates, the inclusion of percent of pasture land does not materially impact the point estimate.

An alternative specification is random-effects. Now, the v_i country-specific term is assumed to be uncorrelated with g_{it}, l_{it}, X_{it}, and W_{it}, while ϵ_{it} is again allowed to be heteroskedastic. The random-effects estimator utilizes the information both within countries as well as across countries, so it is possible to include Z_i in the regressions. Columns (3)

Table 6. Panel Regression Results

	Dep Variable: Log Agricultural Output per Hectare (intl $)						
Exp Variables[a]	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gini coefficient - g_i	−0.49**	−0.50**	−0.50**	−0.49**	−0.42**	−0.48**	−0.63**
	(4.08)	(4.22)	(4.56)	(4.58)	(3.40)	(4.82)	(8.62)
Log avg. farm size - $\ln l_i$	0.02	0.02	0.02	0.02	−0.03	−0.06**	−0.05**
	(1.38)	(1.23)	(1.20)	(1.30)	(1.47)	(2.88)	(2.72)
Inputs							
Log livestock per ha	0.39**	0.39**	0.41**	0.41**	0.25**	0.32**	0.40**
	(9.69)	(9.76)	(11.36)	(11.13)	(8.92)	(12.18)	(17.14)
Log fertilizer per ha	0.04**	0.04**	0.04**	0.04**	0.02**	0.01**	0.01
	(4.04)	(4.13)	(4.11)	(4.03)	(2.53)	(2.27)	(1.22)
Log tractor per ha	0.03**	0.03**	0.03**	0.03**	0.07**	0.07**	0.11**
	(2.60)	(2.85)	(3.31)	(3.15)	(7.12)	(6.93)	(9.32)
Log labor per ha	0.07**	0.07**	0.10**	0.09**	0.36**	0.26**	0.19**
	(2.43)	(2.53)	(3.68)	(3.43)	(11.19)	(7.67)	(6.18)
Land quality							
% Irrigated	1.37**	1.37**	1.28**	1.27**	0.27*	0.34**	0.63**
	(8.84)	8.84)	(8.64)	(8.59)	(1.67)	(2.19)	(5.75)
% Permanent pasture		0.24		−0.17		−0.56**	−0.34**
		(1.07)		(0.98)		(6.37)	(4.65)
Research effort							
Log agric. R&D expend per ha							0.04**
							(5.40)
Constant	3.35**	3.29**	2.42**	2.52**	3.34**	2.94**	3.61**
	(7.92)	(7.48)	(5.99)	(5.86)	(10.44)	(9.51)	(12.93)
Country controls (Z)[b] included	No	No	Yes	Yes	Yes	Yes	Yes
Observations	1,159	1,159	1,159	1,159	1,159	1,159	1,011
Hausman test statistic[c]			5.63	20.43			
Hausman test p-value			0.99	0.99			
Wooldridge test statistic[d]			4.06	4.04			
Wooldridge test p-value			0.05	0.05			
Method	FE	FE	RE	RE	FGLS	FGLS	FGLS
ϵ_{it} autocorrelation	none	none	none	none	AR(1)	AR(1)	AR(1)

Notes: Absolute values of robust t-statistics are given in parentheses. Single astrisk(*) denotes significance at 10%, double astrisk(**) denotes significance at 5%.

[a] All specifications includes total fertility rate, life expectancy, and year dummies.

[b] Z includes the Kraufman, Kraay, and Zoido-Lobaton (2002) index of institutions, dummies for legal origin from La Porta et al. (1999), and land quality from Peterson (1987).

[c] Hausman statistic is distributed as χ^2_{43} in column (3) and χ^2_{44} in column (4).

[d] Wooldridge test statistic is distributed as F(1,53) in both column (3) and column (4).

and (4) report the results of random-effects regressions. As can be seen, the results are nearly identical to the fixed-effects estimates. Hausman test statistics comparing column (3) to column (1) and column (4) to column (2) are reported as well. From these, the hypothesis that differences in the estimated coefficients are purely random cannot be rejected. The random effects specification is preferred due to its efficiency advantages.

Both the fixed-effect and random-effect specifications assumed that ϵ_{it} was heteroskedastic but did not allow for autocorrelation. This assumes that shocks do not persist over time, which may be questionable considering the nature of agricultural production and the interpolation of data. Wooldridge (2002, p. 282) provides a test for serial correlation in panel data and the values of the test statistic are reported in table 6 as part of the random-effects regressions. In both cases, the hypothesis of no serial correlation is rejected.

To control for this, the error structure is modified to follow an AR(1) process. The error term is now $\epsilon_{it} = \rho_i \, \epsilon_{i,t-1} + \eta_{it}$, where ρ_i is between zero and one and η_{it} is an iid noise term. Note that the autoregressive parameter ρ_i is allowed to vary by country. Requiring this term to be constant across countries does not impact the results. With this error structure, the coefficients can be estimated by feasible generalized least squares (FGLS) methods. These results are reported in columns (5) and (6) of table 6. The point estimate for the Gini coefficient falls somewhat in both cases. With the percent pasture land included, though, the estimate is very close to the standard random effects estimates. The point estimate on the percent of pasture land is now significant as well and similar to that found in the OLS specification in column (6) of table 4.

The final column of table 6 presents results estimated by the FGLS technique and assuming an AR(1) error process, but includes a control for R&D expenditure per hectare. This limits the sample to forty-three countries and 1,011 observations. The point estimate on the Gini coefficient rises, as it did under the OLS regressions, and remains significant. In addition, the coefficient on percent pasture land falls. This suggests a positive relationship between agricultural R&D expenditure and both land inequality and the percent pasture land.[6]

[6] Thanks to an anonymous referee for pointing out this implied correlation between inequality and research.

Interestingly, these correlations run counter to the predictions of de Janvry (1978) and de Janvry, Sadoulet, and Fafchamps (1989), who predict that inequitable land distributions depress agricultural research efforts. Further research, though, would be necessary to make any stronger claims regarding the interaction of land inequality and agricultural research.

An interesting departure from the fixed and random-effects results is found in columns (6) and (7), which show that average farm size has a statistically significant negative effect on productivity. This may reflect differences in land quality not captured by the Peterson index. Alternatively, it may also reflect failures in land markets which keep average holding size larger than is optimal. The estimate is such that the difference between the median Latin American holding size of 17.70 hectares and the median South Asian holding size of 2.32 hectares is associated with a productivity difference of 13%.

The values of the coefficients on the input variables are in line with previous work. The coefficients on fertilizer, tractors, and labor are very close to previous work (Craig, Pardey, and Roseboom 1997), while the coefficient on livestock is higher than some previous studies (Fulginiti and Perrin 1993; Frisvold and Ingram 1995).

Overall, the results of the panel regressions confirm the general findings of the OLS results, while showing a smaller absolute size of the point estimate on the Gini coefficient. Presuming that the set of other control variables is comprehensive, then the fact that the Gini coefficient is significant and negatively related to aggregate productivity implies that there is heterogeneity of productivity by farm size within countries.

It is interesting to consider what these results imply about the degree of heterogeneity of productivity. Consider for the moment that there is an identical IFSP relationship within every country, so that production on farm i is $\ln Y_i = \alpha + \delta \ln L_i$, where Y_i is output, and L_i is farm size. The estimated effect of the Gini coefficient contains information on the size of δ. If $\delta = 1$, then the distribution of land should not matter in aggregate productivity. The smaller is δ, the more dramatic the effect of the Gini coefficient on aggregate productivity. Making some simplifying assumptions about the nature of the Lorenz curve, it is possible to back out a value for δ from the estimated cross-country regressions. From table 6, column (6), the estimated coefficient on the Gini is −0.48. This estimate implies a value of $\delta = 0.807$ and Appendix B describes in detail how this is calculated.

This value is similar to those found by micro-level studies of the IFSP. The value of δ ranges between 0.58 and 0.91 for Indian states considered by Bhalla and Roy (1988), while Carter (1984) finds δ to fall between 0.58 and 0.66 depending on specification. Benjamin (1995) finds $\delta = 0.82$ in his initial fixed-effects specification and $\delta = 0.78$ in initial random-effects specifications. Lamb (2003) finds δ of 0.89 in random-effects regressions for household profits. Results using total labor hours rather than profits show δ ranging between 0.80 to 0.92.

However, both Benjamin (1995) and Lamb (2003) find that $\delta = 1$ once they control for unobserved heterogeneity in land quality by farm. This suggests an important caveat to the results in this article. While the regressions here do control for average land quality by country, they do not control for heterogeneity of land quality within countries. Specifications that do control for heterogeneity may show the results here to be spurious, similar to the dissappearance of the IFSP in the work of Benjamin and Lamb. A suitable measure of heterogeneity in land quality does not appear to be available at this point, so further research will be necessary to address this issue completely.

Conclusion

This article has quantified the effect of land distribution on cross-country agricultural productivity by using data from Deininger and Squire (1998) regarding the Gini coefficient for the size of operational land holdings within countries. The empirical work shows a significant negative relationship between the Gini coefficient and output per hectare. This effect persists even after controlling for input use, land quality, human capital, institutions, and agricultural research effort. These results support previous work on the advantages of unimodal or broad-based distributions of land. The negative relationship of land inequality and productivity implies heterogeneity in productivity by farm size, a finding consistent with the literature on the inverse farm-size productivity relationship.

Point estimates imply that a drop in the Gini coefficient of one standard deviation would increase output per hectare by 8.5%. This effect is net of any changes in input use or land utilization that may occur with a change in land distribution, so the total effect of a drop in the Gini coefficient may in fact lead to an even larger gain in productivity.

The fact that such a productivity effect exists suggests that the distribution of land within countries is not optimal. This implies that land markets are not functioning properly, something which has been well documented and discussed frequently in micro-level research. The cross-country analysis shows that these imperfections have macro-level consequences. In addition, it offers the possibility of estimating the potential benefits of land market reforms that allow operational holdings to be more equitably distributed.

[Received August 2005; accepted April 2006.]

References

Banerjee, A., and L. Iyer. 2005. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review* 95:1190–213.

Bardhan, P.K. 1973. "Size, Productivity, and Returns to Scale: An Analysis of Farm-Level Data in Indian Agriculture." *Journal of Political Economy* 81:1370–86.

Barrett, C.B. 1996. "On Price Risk and the Inverse Farm Size-Productivity Relationship." *Journal of Development Economics* 51:193–215.

Benjamin, D. 1995. "Can Unobserved Land Quality Explain the Inverse Productivity Puzzle?" *Journal of Development Economics* 46:51–84.

Berry, A.R., and W.R. Cline. 1979. *Agrarian Structure and Productivity and Developing Countries*. Baltimore: The Johns Hopkins University Press.

Besley, T., and R. Burgess. 2000. "Land Reform, Poverty Reduction, and Growth: Evidence from India." *Quarterly Journal of Economics* 115:389–430.

Bhalla, S.S. 1988. "Does Land Quality Matter?" *Journal of Development Economics* 29:45–62.

Bhalla, S.S., and P. Roy. 1988. "Mis-Specification in Farm Productivity Analysis: The Role of Land Quality." *Oxford Economic Papers* 40:55–73.

Binswanger, H.P., and K. Deininger. 1997. "Explaining Agricultural and Agrarian Policies in Developing Countries." *Journal of Economic Literature* 35:1958–2005.

Binswanger, H.P., K. Deininger, and G. Feder. 1995. "Power, Distortions, Revolt and Reform in Agricultural Land Relations." In J. Behrman, and T.N. Srinivasan, eds. *Handbook of Development Economic, Vol. III*. pp. 2659–772, Amsterdam: Elsevier.

Carter, M.R. 1984. "Identification of the Inverse Relationship between Farm Size and Productivity: An Empirical Analysis of Peasant Agricultural Production." *Oxford Economic Papers* 36:131–45.

Craig, B.J., P.G. Pardey, and J. Roseboom. 1997. "International Productivity Patterns: Accounting for Input Quality, Infrastructure, and Research." *American Journal of Agricultural Economics* 79:1064–76.

de Janvry, A. 1978. "Social Structure and Biased Technical Change in Argentine Agriculture." In H.P. Binswanger and V.W. Ruttan, eds. *Induced Innovation: Technology, Institutions, and Development*. Baltimore, MD: The Johns Hopkins University Press, pp. 297–323.

de Janvry, A., and E. Sadoulet. 1989. "A Study in Resistance to Institutional Change: The Lost Game of Latin American Land Reform." *World Development* 17:1397–1407.

de Janvry, A., E. Sadoulet, and M. Fafchamps. 1989. "Agrarian Structure, Technological Innovations, and the State." In P. Bardhan, ed. *The Economic Theory of Agrarian Institutions*. Oxford: Clarendon Press, pp. 356–82.

Deininger, K., and L. Squire. 1998. "New Ways of Looking at Old Issues: Inequality and Growth." *Journal of Development Economics* 57:259–87.

Diaz, A. 2000. "On the Political Economy of Latin American Land Reform." *Review of Economic Dynamics* 3:551–71.

Engerman, S.L., and K.L. Sokoloff. 1997. "Factor Endowments, Institutions and Differential Paths of Growth Among New World Economies." In S.H. Haber, ed. *How Latin America Fell Behind*. Stanford CA: Stanford University Press, pp. 260–304.

Eswaran, M., and A. Kotwal. 1986. "Access to Capital as a Determinant of the Organization of Production and Resource Allocation in an Agrarian Economy." *Economic Journal* 96:482–98.

Feder, G. 1985. "The Relation Between Farm Size and Farm Productivity." *Journal of Development Economics* 18:297–313.

Frisvold, G.B. 1994. "Does Supervision Matter? Some Hypothesis Tests Using Indian Farm Level Data." *Journal of Development Economics* 43:217–38.

Frisvold, G., and K. Ingram. 1995. "Sources of Agricultural Productivity Growth and Stagnation in Sub-Saharan Africa." *Agricultural Economics* 13:51–61.

Fulginiti, L.E., and R.K. Perrin. 1993. "Prices and Productivity in Agriculture." *Review of Economics and Statistics* 75:471–82.

Hayami, Y., and V.W. Ruttan. 1970. "Agricultural Productivity Differences among Countries." *American Economic Review* 60:895–911.

Hayami, Y., and V.W. Ruttan. 1985. *Agricultural Development: An International Perspective*. Baltimore: Johns Hopkins Univ. Press.

Heltberg, R. 1998. "Rural Market Imperfections and the Farm Size-Productivity Relationship: Evidence from Pakistan." *World Development* 26:1807–26.

Horowitz, A.W. 1993. "Time Paths of Land Reform: A Theoretical Model of Reform Dynamics." *American Economic Review* 83:1003–10.

Jeon, Y., and Y. Kim. 2000. "Land Reform, Income Redistribution, and Agricultural Production in Korea." *Economic Development and Cultural Change* 48:253–68.

Johnston, B.F., and P. Kilby. 1975. *Agriculture and Structural Transformation: Economic Strategies in Late-Developing Countries*. London: Oxford University Press.

Johnston, B.F., and W.C. Clark. 1982. *Redesigning Rural Development: A Strategic Perspective*. Baltimore: The Johns Hopkins University Press.

Kaufmann, D., A. Kraay, and P. Zoido-Lobaton. 2002. "Governance Matters II: Updated Indicators for 2000/01." World Bank Policy Research Department Working Paper.

Kawagoe, T., Y. Hayami, and V. Ruttan. 1985. "The Intercountry Agricultural Production Function and Productivity Differences among Countries." *Journal of Development Economics* 19:113–32.

La Porta, R., F. Lopez de Silanes, A. Schleifer, and R. Vishny. 1999. "The Quality of Government." *Journal of Law, Economics, and Organization* 15:222–79.

Lamb, R.L. 2003. "Inverse Productivity: Land Quality, Labor Markets, and Measurement Error." *Journal of Development Economics* 71:71–95.

Lau, L.J., and P.A. Yotopoulos. 1988. "The Meta-Production Function Approach to Technological Change in World Agriculture." *Journal of Development Economics* 31:241–69.

Masters, W.A. 2003. "Climate, Agriculture, and Economic Development." In K. Wiebe, eds. *Land Quality, Agricultural Productivity, and Food Security*. Northampton, MA: Edward Elgar Publishing, pp. 166–83.

Mundlak, Y. 2000. *Agriculture and Economic Growth: Theory and Measurement*. Cambridge, MA: Harvard University Press.

Otsuka, K., H. Chuma, and Y. Hayami. 1992. "Land and Labor Contracts in Agrarian Economies: Theories and Facts." *Journal of Economic Literature* 30:1965–2018.

Pardey, P.G., and J. Roseboom. 1989. *ISNAR Agricultural Research Indicator Series*. Cambridge, UK: Cambridge University Press.

Peterson, W. 1987. "International Land Quality Indexes." University of Minnesota Dept. of Agricultural and Applied Economics Staff Paper P87-10.

Sokoloff, K.L., and S.L. Engerman. 2000. "Institutions, Factor Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives* 14(1):217–32.

Tomich, T.P., P. Kilby, and B.F. Johnston. 1995. *Transforming Agrarian Economies*. Ithaca, NY: Cornell University Press.

United Nations, Food and Agriculture Organization. 1997. "Report on the 1990 World Census of Agriculture." FAO Statistical Development Series 9.

United Nations, Food and Agriculture Organization. 1999. "FAOSTAT Database." Available at http://faostat.fao.org/default.aspx.

Wiebe, K., M. Soule, C. Narrod, and V. Breneman. 2003. "Resource Quality and Agricultural Productivity: A Multi-country Comparison." In K. Wiebe, ed. *Land Quality, Agricultural Productivity, and Food Security*. pp. 147–65. Northampton, MA: Edward Elger Publishing.

Wooldridge, J.M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

World Bank. 2003. "World Development Indicators." Available at http://devdata.worldbank.org/dataonline/.

Appendix A: Data Sources

The OLS sample is driven by the availability of the Gini coefficient. In some cases, the other explanatory variables are not available in the year in which a Gini coefficient was observed. If possible, the value of the other explanatory variables was interpolated from existing data. The sections below describe the source of the variables as well as the interpolation technique used, if applicable. The final section of the Appendix describes the interpolation techniques used to obtain the larger panel data set.

Agricultural Output and Inputs

Total *output* is obtained from the FAOSTAT (FAO 1997) database and is the total value of all agricultural production after deductions for feed and seed. This value is a price-weighted sum of the quantity of all agricultural outputs given in terms of international dollars. The international dollar was developed by the FAO to avoid having to use market exchange rates to compare the value of output across countries. It is derived from the Geary-Khamis formula that calculates simultaneously the relative price of each component of output and the implicit exchange rate of each country's currency with respect to the international dollar.

Data on inputs are from the FAOSTAT database as well. The measure of *land* is the total hectares of agricultural land, which consists of arable land, permanent crop, land and permanent pasture land. *Livestock* is the number of cow equivalents, a measure commonly used in the cross-country literature. It is calculated using weights obtained from Hayami and Ruttan (1985). The weighting is as follows: 1 horse = 1 mule = 1 buffalo = 1.25 cattle = 1.25 asses = 0.9 camels = 5 pigs = 10 sheep = 10 goats = 100 chickens = 100 ducks = 100 geese = 100 turkeys. An alternative method of weighting livestock use in agriculture is suggested by Craig, Pardey, and Roseboom (1997) and divides livestock into those animals used primarily for traction power and those used for breeding purposes. Using the Craig, Pardey, and Roseboom (1997) livestock measures does not materially impact the results of the regressions.

Tractors is measured by the number of agricultural tractors in use and are all assumed to be 30 horsepower. This measure excludes two-wheeled tractors and garden tractors and so is not a perfect measure of capital services available. However, this is the only data set on capital services in agricultural that covers a wide range of countries and time periods. *Fertilizer* is the total metric tons used of nitrogenous, phosphate, and potash fertilizer. *Labor* is measured as the total economically active population in agriculture.

The FAOSTAT database only goes back to 1961. There are thirty-one observations of the Gini coefficient and average farm size that occur prior to this and they go back as far as 1958. For these years the value of output, land, livestock, tractors, fertilizer, and labor is extrapolated from the existing data. The method is to calculate the growth rate of each variable over the five years 1961-65. It is then assumed that this growth rate obtained for the years prior to 1961 and the variables are extrapolated by applying this growth rate backwards from 1961.

Land Quality

The *land quality index* of Peterson (1987) is determined by the predicted value of agricultural land per acre in a country divided by the average value of land per acre across all countries, which in Peterson's work consisted of 126 countries. The log of the value of agricultural land per acre is a weighted linear function of (a) nonirrigated crop land as a percent of all agricultural land, (b) irrigated land as a percent of all arable crop land, and (c) the log of the long-run average annual precipitation. The weights on each of the three elements were derived from farm level data in the United States. There is no way to interpolate these data, so any country which did not have an observation in the Peterson data was excluded.

Another land quality index is available from Wiebe et al. (2003) that is based on spatially referenced data within countries on soil types and precipitation. This index covers a smaller sample of countries and using it in place of or alongside of the Peterson index does not materially impact the results reported in this article.

The *percent of land irrigated* and *percent of land in pasture* are derived as the amount of irrigated land and amount of permanent pasture land, respectively, divided by total agricultural land. These data are obtained from the FAOSTAT database. For the same thirty-one observations cited in the previous section, a similar interpolation method was used to obtain values that occur prior to 1961.

Human Capital

Both *life expectancy* and *total fertility rate* are obtained from the World Bank (2003). Missing values were interpolated on the assumption that these variables follow relatively smooth paths over time. For total fertility rate, the technique is as follows. The value at time s which falls between actual observations at time t and $t + n$ is calculated as $\ln TFR_s = \ln TFR_t + (s - t)\ (\ln TFR_{t+n} - \ln TFR_t)/n$. For life expectancy, there are similar gaps in the data and an identical interpolation equation is used.

Institutional Quality

The *institutional quality index* is derived by Kaufmann, Kraay, and Zoido-Lobaton (2002). Their index aggregates several components: voice and accountability, political stability, government effectiveness, regulatory quality, rule of law, and control of corruption. The value for an individual country averages the value of each of the six components. The components themselves are each measured relative to the worldwide average which is set to zero. Dummy variables for *legal origin* are obtained from La Porta et al. (1999). This characterizes the basic legal structure in each country as coming from either a British, French, Socialist, Scandanavian, or German background.

Neither the institutional quality index nor legal origins can be interpolated, so any country without observations of these variables was eliminated from the sample.

Agricultural Research Effort

Data on *agricultural R&D expenditures* are obtained from Pardey and Roseboom (1989) and reflects the expenditures made by national agricultural research systems, measured in PPP terms. It thus excludes private R&D efforts, although the authors note these are generally minimal except in the most advanced countries. Data are available on a limited basis between the years 1960 and 1986; so in some cases, interpolation of expenditure data was made. The method was to calculate expenditure as $\ln RDX_s = \ln RDX_t + (s - t)\ (\ln RDX_{t+n} - \ln RDX_t)/n$, where s is the time period interpolated, and t and $t + n$ are time periods in which research expenditure was observed.

Due to the nature of research and development, the more desirable measure would be lagged values or an aggregation of lagged values of research expenditures. The lack of data limits this possibility and the current research effort is assumed to be a good proxy for overall research effort.

Panel Data Sources

The sources of the data elements in the panel regressions are identical to the OLS regressions. The difference lies in the size of the sample. For the panel data, both the Gini coefficient and the average farm size were interpolated to create a data set with a longer time series dimension. There are fifty-four countries which have at least two actual observations of the Gini coefficient and average farm size. For the years lying between these observations values were interpolated. For the Gini coefficient, this was done as simple linear interpolation. The Gini at time s was calculated as $g_s = g_t + (s - t)\ (g_{t+n} - g_t)/n$, where g_{t+n} and g_t are the observed values. This was then done for every year lying between t and $t + n$. For the average farm size, a constant growth rate among observations was assumed. This led to an interpolation of the following form, $\ln l_s = \ln l_t + (s - t)\ (\ln l_{t+n} - \ln l_t)/n$. All other data were either obtained directly or interpolated for the missing years through the techniques described in the preceding sections.

Appendix B: Implied Farm Size Elasticity of Productivity

The point estimate on the Gini coefficient can be used to solve for the implied elasticity of productivity with respect to farm size. The Gini coefficient is based on the Lorenz curve, and this curve can be modelled as $s_L = s_F^{\beta}$, where $s_L \in [0, 1]$ is the share of total land, $s_F \in [0, 1]$ is the share of farms, and β is a parameter that represents the inequality of the distribution. β is equal to or greater than one, with $\beta = 1$ indicating perfect equality. The Gini coefficient, g, can be written as $g = \frac{\beta-1}{\beta+1}$ and this will be used to back out the parameter β from the data.

With no loss of generality, set the average farm size in the whole country to be equal to one. The average farm size of the farm at the nth percentile is then simply $\frac{\partial s_L}{\partial s_F} = \beta n^{\beta-1}$. In reduced form, the output of any farm i is $y_i = Al_i^{\delta}$, where y_i is output, A is a productivity term fixed for all farms, and l_i is the size of farm i. The parameter δ represents the elasticity of output with respect to farm size. If the IFSP relationship holds, this will be less than one. For the farm at the nth percentile, output is $y_n = A(\beta n^{\beta-1})^{\delta}$.

Aggregating over all farms, total output in the economy is therefore

$$\text{(B.1)} \quad Y = \int_0^1 A(\beta n^{\beta-1})^{\delta} dn$$

which can be reduced to

$$\text{(B.2)} \quad Y = A\frac{\beta^{\delta}}{(\beta-1)\delta+1}.$$

Dividing both sides by the total amount of land L and taking logs gives an expression in terms of log output per hectare

$$\text{(B.3)} \quad \ln\frac{Y}{L} = \ln\frac{A}{L} + \delta\ln\beta - \ln((\beta-1)\delta+1).$$

From, (7), it can be seen that if there is no IFSP relationship ($\delta = 1$); then output per hectare is simply $\ln Y/L = \ln A/L$, or the distribution of land has no effect on output. If there is no inequality ($\beta = 1$), then output is again $\ln Y/L = \ln A/L$ and the degree to which there is an IFSP relationship is irrelevant because farms do not vary in size.

Equation (7) can be used to establish the size of δ given the response of output per hectare to changes in land distribution. Using the estimated production functions, we know how $\ln Y/L$ responds to a change in the Gini coefficient. Consider a small change in the Gini coefficient, dropping from the panel sample mean value of 0.65 to 0.64, or $\Delta g = -0.01$. This change in the Gini also implies that β fell from 4.71 to 4.56, a change of -0.15. Given the point estimate in table 6, column, (6) the fall in the Gini coefficient also implies an increase in log output per hectare of 0.0048. Therefore, we have $\frac{\Delta \ln Y/L}{\Delta\beta} = \frac{0.0048}{-0.15} = -0.032$.

Using equation (7), we can now ask what value of δ is consistent with $\frac{\Delta \ln Y/L}{\Delta\beta} = -0.032$, holding $\ln A/L$ constant. Values of δ were iterated through in (7) until one was found such that $\Delta\beta = -0.15$ led to $\Delta \ln Y/L = 0.0048$. This value was found to be $\delta = 0.807$.