



Inverse Productivity or Inverse Efficiency? Evidence from Mexico

Justin Kagin, J. Edward Taylor & Antonio Yúnez-Naude

To cite this article: Justin Kagin, J. Edward Taylor & Antonio Yúnez-Naude (2016) Inverse Productivity or Inverse Efficiency? Evidence from Mexico, The Journal of Development Studies, 52:3, 396-411, DOI: [10.1080/00220388.2015.1041515](https://doi.org/10.1080/00220388.2015.1041515)

To link to this article: <https://doi.org/10.1080/00220388.2015.1041515>

 [View supplementary material](#) 

 Published online: 13 Aug 2015.

 [Submit your article to this journal](#) 

 Article views: 494

 [View Crossmark data](#) 

 Citing articles: 5 [View citing articles](#) 

Inverse Productivity or Inverse Efficiency? Evidence from Mexico

JUSTIN KAGIN^{*,**}, J. EDWARD TAYLOR^{*} & ANTONIO YÚNEZ-NAUDE[†]

^{*}University of California, Department of Agricultural and Resource Economics, Davis, CA, USA, ^{**}Kagin's Consulting, Vacaville, CA, USA, [†]El Colegio de México, Mexico City, Distrito Federal, Mexico

(Final version received March 2015; final version accepted March 2015)

ABSTRACT *Using a unique panel data set from rural Mexico, we find strong evidence of a negative relationship between farm size and both productivity and technical efficiency: large farms not only have a lower value of output per hectare than small farms, they also produce further from the efficiency frontier. Our findings suggest that, in spite of the ongoing transformation of agricultural supply chains and economists' recommendations for small farmers to exit crop production, there may be sustained advantages for smallholder farms. Our analysis offers new insights into inverse-farm size relationship, the productivity–efficiency relationship, and the use of stochastic frontier techniques.*

1. Introduction

Since Chayanov ([1926] 1991), the received wisdom in development economics has been that there is an inverse relationship between farm size and land productivity, implying that small and medium-size farms produce crops with higher yields than large holders. This would appear to be puzzlingly inefficient, because under the conventional assumptions of constant returns to scale and perfect competition, land productivity should be equalised across producing units.¹ Researchers have advanced a number of different explanations for the inverse relationship, including risk (Barrett, 1996), labour market dualism and other market failures (Eswaran & Kotwal, 1986; Heltberg, 1998; Sen, 1966), and omitted variable bias (Benjamin, 1995; Bhalla & Roy, 1988; Carter, 1984).²

The dynamics of small-farm efficiency have taken on a new research imperative in light of ongoing transformation of agricultural markets and supply chains. Pingali, Khwaja, and Meijer (2005, p. 61) point out threats to the survival of smallholders resulting from the concentration of food trade. Collier (2008, p. 72) offers a particularly grim assessment:

The peasant life forces millions of ordinary people into the role of entrepreneur, a role for which most are ill suited [...] their mode of production is ill suited to modern agricultural production.

Technical efficiency is vital to the prospects of small-farm agriculture; however, despite a plethora of efforts to document and explain the inverse productivity relationship, the implications for efficiency are not well understood. While a number of empirical studies provide evidence that an inverse productivity relationship exists, none to our knowledge has examined changes in either

Correspondence Address: Justin Kagin, 326 Cooper School Rd. Vacaville, CA. 95687, Email: justin.kagin@gmail.com

An Online Appendix is available for this article which can be accessed via the online version of this journal available at <http://dx.doi.org/10.1080/00220388.2015.1041515>

this or the efficiency of small versus large farms over time. We are not aware of any study that has seriously explored the implications of the inverse productivity relationship for technical efficiency or vice versa, despite pointing to the need and relevance of doing so (for example, Heltberg, 1998, p. 1812).

This paper uses a unique panel data set from rural Mexico to test the implications of an inverse productivity relationship for efficiency. We first present a simple illustration of the difference between inverse productivity and inverse efficiency. We then use the Mexico panel to test whether the inverse productivity relationship holds. We find strong evidence of a negative relationship between farm size and land productivity, even while controlling for farm fixed effects. Whether this reflects higher efficiency on small farms is an empirical question, given conflicting theories about the causes of the inverse productivity relationship. By linking agricultural productivity with technical efficiency in a stochastic frontier analysis (SFA), we test whether production efficiency and the efficiency frontier itself vary with farm size in ways consistent with an inverse efficiency relationship. Our analysis reveals that larger farms have a lower value of output per hectare, and they operate further from the efficiency frontier, on average, than do small farms.

Our data make it possible to control for input quality and human capital, inputs of the production function usually excluded from farm SFA models. They also make it possible, in the inefficiency component of the SFA estimation, to control for variables likely to influence inefficiency. That incentives for efficiency increase with market competition is a basic tenet of the literature on trade, and it has been exploited in a variety of models (Baily & Gersbach, 1995; Green & Mayes, 1991; Griffiths, 2001). High transaction costs isolate farmers from outside markets in much the same way as import restrictions protect local producers, with well-known negative implications for production efficiency. Membership of an indigenous community may have a similar effect, isolating producers from markets as well as information, while also constraining farmers to conform to traditional standards or norms that may or may not be efficient in a broader market (that is, regional or national). Research on the economics of identity in excluded groups (Akerloff & Kranton, 2000; Fang & Loury, 2005; Oxoby, 2004) employs models with multiple economic equilibria; individuals from discriminated and segregated populations rationally choose norms with low economic outcomes due to their high social-psychological value. Traditional farming methods may therefore be optimal within an indigenous community but not necessarily efficient when compared with farmers outside the group. International labour migrants provide knowledge as well as remittances to their countries of origin, and this has spawned several studies on knowledge transfer and its productivity implications in the migration literature (Boucher, Stark, & Taylor, 2009; Docquier, Faye, & Pestieau, 2008; Docquier & Marfouk, 2006).

Including these variables in our model enables us to test the robustness of our findings on inverse efficiency. We find that the inverse productivity and efficiency relationships are robust to the inclusion of controls for input quality and farmer human capital in the production function; however, the inverse efficiency relationship disappears when we control for market access, ethnicity, and international migration in the frontier regression.

2. Productivity and Efficiency

Research on the inverse relationship tends not to distinguish between productivity and efficiency; indeed, many studies use the two terms interchangeably. However, evidence of an inverse relationship for one does not necessarily imply an inverse relationship for the other. Figure 1 represents a simple output isoquant with land and labour inputs. According to the inverse farm size relationship, farmers with a large amount of land (T_L) produce less per hectare than small farms (T_S), as shown on the graph. This might be the result of market imperfections creating a deviation in relative input prices between large and small farms, implicit in most explanations for the inverse relationship. Yet, if Q represents the efficient frontier, both the large and small farms in this diagram produce at complete technical efficiency, the minimum combination of inputs given output.³

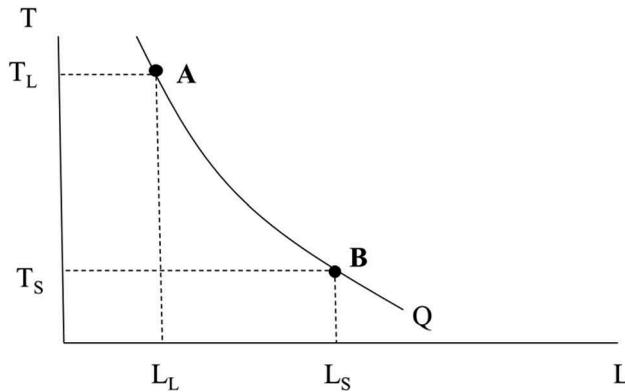


Figure 1. Technically efficient farmers.

Note: T is land used in crop production where T_L is land used for a large farmer and T_S is land used for a small farmer. L is labor used where L_L is labour used for a large farmer and L_S is labour used for a small farmer. Q is the efficient frontier output.

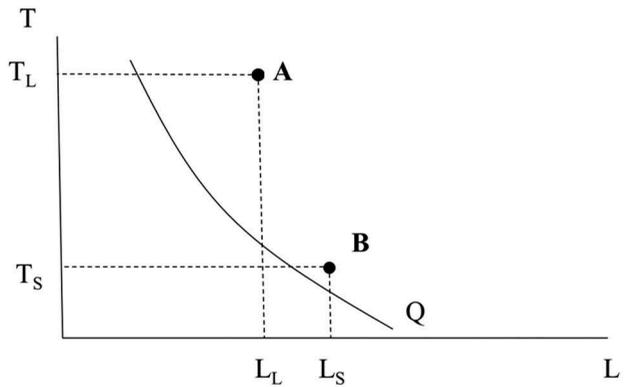


Figure 2. Inverse technical efficiency.

Note: T is land used in crop production where T_L is land used for a large farmer and T_S is land used for a small farmer. L is labor used where L_L is labour used for a large farmer and L_S is labour used for a small farmer. Q is the efficient frontier output.

Figure 2 illustrates the case of an inverse relationship with respect to both land productivity and efficiency: small farms produce closer to the efficiency frontier than large ones. A result like this would indicate that small farms use inputs in a more technically efficient manner than large farms. Over time, as technology changes, there may be shifts in the frontier, which may affect the relative efficiency of small and large farmers. Figure 3 depicts Collier’s argument that technology changes have benefited large farmers over small ones, increasing their relative efficiency. A shift in the isoquant from Q_1 to Q_1' , which takes large farms (A) along with it while leaving small farms (B) behind, appears to increase the inefficiency of small farms.

In special cases inverse productivity may imply inverse efficiency. Sen (1966) argued that, although large holders may be able to take better advantage of certain types of technology and have economies of scale in production, small and medium producers have greater access to family labour relative to their farm size. As farm size increases, so does the reliance on wage labour, which must be monitored. In such a case, higher land productivity and intensity of labour use could be associated with greater efficiency on small farms: to achieve the same output level, large farms would have to utilise more labour days than small farms (or else incur monitoring costs). On the other hand, if smallholder households cannot sell all of the labour they want, so that family workers are ‘trapped’ on the farm,

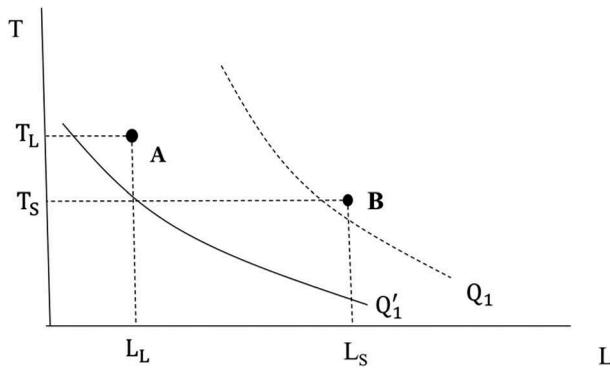


Figure 3. Technology changes increase relative productivity and technical efficiency of large farms.

Note: T is land used in crop production where T_L is land used for a large farmer and T_S is land used for a small farmer. L is labor used where L_L is labour used for a large farmer and L_S is labour used for a small farmer. Q_1 is the initial efficient frontier output, Q_1' is the efficient frontier after the technology change.

this could result in high production per hectare. Whether or not this is technically efficient depends on whether the household could achieve a higher output with the land and labour at its disposal. In short, an agricultural sector could be economically inefficient (with different farmers using different decision prices), yet individual farms may be technically efficient (minimising input combinations to achieve a given output level).

A number of studies report evidence of an inverse relationship between farm size and productivity. Recently, some studies have questioned the relative efficiency of small farms, particularly in the context of expanding global markets for agricultural products and the transformation of the agricultural supply chain. The exigencies of modern markets, including new requirements for producers to meet grades and standards in product markets, create a variety of transaction costs in input and output markets that are likely to favour large producers. Even if small farmers can manage their labour more efficiently, it is possible that scale economies in the utilisation of skills and technology, access to capital, and the organisation of logistics and trading have come to offset smallholder advantages. According to Barrett, Reardon, and Webb (2001 p. 321), 'Missing credit markets can impede diversification into activities or assets characterized by substantial barriers to entry'. If access to new technologies shifts the efficiency frontier inward for large farmers, leaving small farmers behind, it could result in a positive relationship between measured efficiency and farm size. On the other hand, access to off-farm income, including migrant remittances, might loosen liquidity and risk constraints, facilitating technology adoption on small farms, with implications for efficiency (Pfeiffer, Lopez-Feldman, & Taylor, 2009).

Smallholders' technical efficiency has ramifications for their potential role in combating poverty and enhancing food security. If scale effects are important in today's market environment, smallholder agriculture may not be capable of sustained economic growth. Targeting small-scale agriculture may not be justified from a production perspective if this sector is inherently inefficient. On the other hand, if small farms boast not only productivity but also efficiency advantages, it could be that a heterogeneous farm structure, in which small farms coexist with large ones, is consistent with promoting agricultural growth. As Figures 1–3 illustrate, whether or not small farms have a productivity and/or technical efficiency advantage over large farms, and whether they are able to hold onto their productivity advantage and remain competitive as agricultural markets evolve, are empirical questions.

3. Econometric Models, Data, and Findings

Most empirical investigations into the inverse productivity relationship use descriptive or cross-section analysis of differences in output, output value, or profits per unit of land area. These approaches are

not likely to provide reliable tests of an inverse relationship with respect to productivity or efficiency. The first does not control for observed differences among farms (for example soil quality), and the second does not control for the confounding effects of unobserved variables (for example farmer skill). Only a few studies of the inverse productivity relationship use panel data and fixed effects methods to control for unobservables at the farm level. None to our knowledge consider both productivity and efficiency using panel methods.

The key data input for our analysis are from the 2003 and 2008 rounds of the Mexico National Rural Household Survey (*Encuesta Nacional de Hogares Rurales de Mexico*; Spanish acronym ENHRUM). The ENHRUM provides a matched longitudinal data set on assets, socio-demographic characteristics, production, income sources, and migration from a nationally representative sample of rural households first surveyed in January and February 2003. The 2003 (2002 data) sample includes 1,782 households in 14 Mexican states; of these, 1,543 were successfully resurveyed in 2008 (2007 data).

el Instituto Nacional de Estadística, Geografía e Informática (INEGI), Mexico's national information and census office, designed the sampling frame to provide a statistically reliable characterisation of Mexico's population living in rural areas, or communities with fewer than 2,500 inhabitants. For reasons of cost and tractability, individuals in hamlets or disperse populations with fewer than 500 inhabitants were not included in the survey. The result is a sample representative of more than 80 per cent of the population that the Mexican government considers to be rural. Round II of the ENHRUM, carried out in 2008, revisited the 2003 sample of households. In both rounds, input, output, and time use data were gathered on all household crop and non-crop activities in the year prior to the survey (2002 and 2007). The ENHRUM is unique in providing detailed production data from a household panel covering two years over a five-year period (2002–2007). To our knowledge, this analysis is the first to identify the determinants of productivity and efficiency in Mexico's agricultural economy over time using household panel data.

Table 1 summarises the variables used in our econometric models. It also splits the sample in half by farm size and compares variable means (the sample splits almost evenly at a farm size of three hectares). The descriptive statistics provide some informal support for the inverse farm size relationship: the value of agricultural output per hectare is significantly higher on small than large farms.

Figure 4 shows the distribution of our data, comparing the logged value of operated land to logged value of output per hectare. The graph also provides preliminary support for the inverse farm size relationship; despite variation, there is a clear downward trend in the productivity of land as operated land size increases.⁴

Table 1. Summary statistics

Average Variable	Small (≤ 3 hectares)	Large (> 3 hectares)	Difference
Average output per hectare	10,047.10	4,575.56	**
Average land (hectares)	1.50	9.89	***
Average labour (hours)	84.86	144.07	***
Average capital (MXN)	606.48	2134.71	***
Average purchased inputs (MXN)	1866.16	6,702.23	***
Household head education	3.74	4.04	*
Good land (% of total)	35.60	34.70	
Flat land (% of total)	44.80	59.80	***
Irrigated land (hectares)	0.29	2.01	***
Only dirt roads (% of total)	12.18	6.43	***
Household head speaks an indigenous language (% of total)	38.00	24.00	***
Number of households	788	575	

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

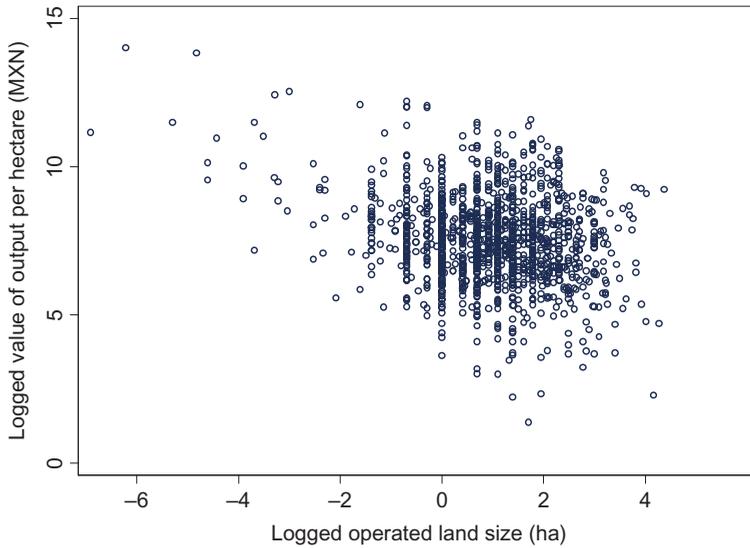


Figure 4. Operated land size and output per hectare of land.

Testing the Inverse Relationship: A Panel Approach

The traditional approach to test the inverse relationship is to regress farm output per hectare of operated land on farm size or cultivated land area. A negative coefficient on land indicates an inverse relationship (Heltberg, 1998). Such a simple regression, however, does not account for bias from unobserved farmer heterogeneity or other variables, which can easily confound productivity tests. We used the Mexican panel to estimate the following model to test for the inverse relationship:

$$\ln\left(\frac{Y_{it}}{T_{it}}\right) = \alpha_i + \beta_1 \ln(T_{it}) + \beta_2 \ln\left(\frac{L_{it}}{T_{it}}\right) + \beta_3 \ln\left(\frac{K_{it}}{T_{it}}\right) + \beta_4 \ln\left(\frac{C_{it}}{T_{it}}\right) + \beta_5 \text{year} + \epsilon_{it}. \quad (1)$$

The left-hand variable is the log of agricultural output value per hectare of cultivated land. The right-hand variables include farm size, T_{it} , farm labour per hectare, $\frac{L_{it}}{T_{it}}$ (family and hired labour days), capital per hectare, $\frac{K_{it}}{T_{it}}$ (animals and machinery costs), other input costs, $\frac{C_{it}}{T_{it}}$ (fertiliser, pesticides, and seeds), and the year (1 if the year is 2007); α_i denotes farm fixed effects.⁵ In a constant-returns-to-scale economy with perfect factor markets, there should be no observed differences in productivity across farm sizes; that is, $\beta_1 = 0$.

Table 2 reports the results from both a random effects (Column A) and fixed effects (Column B) regression for land productivity. The estimated coefficients represent the percentage changes in the dependent variable associated with a 1 per cent increase in each of the explanatory variables. The results support the inverse farm size relationship. The random effects model finds that, other things being equal, a 1 per cent increase in farm size is associated with an insignificant 0.06 per cent reduction in output value per hectare. The fixed effects model shows a much greater and significant negative effect of -0.25 per cent. Farm fixed effects control for time invariant farm-household variables, observed or unobserved, including soil quality (which changed little between the survey years) and farmer management expertise. The contrast in model results reveals the consequence of not controlling for these effects: an underestimation of the inverse farm size relationship.

These findings offer compelling support for the inverse relationship in Mexican agriculture, mirroring findings from panel studies conducted in other countries. They suggest a sustained

Table 2. Random and fixed effects estimates of land productivity

Variable	Value of crop output (MXN) per hectare	
	RE	FE
Ln (average Land)	-0.055 (0.038)	-0.226*** (0.071)
Ln (average labour/average land)	0.195*** (0.030)	0.197*** (0.046)
Ln (average capital/average land)	0.043*** (0.010)	-0.004 (0.018)
Ln (average other costs/average land)	0.156*** (0.015)	0.042* (0.023)
Year	0.168*** (0.059)	0.144** (0.065)
Constant	5.98*** (0.15)	6.86*** (0.21)
Observations	1,361	1,361
Number of households	842	842

Notes: Ln is the natural logarithm, RE is random effects, FE is fixed effects. Standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

advantage for small farms in terms of land productivity over time; however, they do not shed light on differences in total factor productivity or the relative technical efficiency of small farms.

Testing the Inverse Efficiency Hypothesis

Measuring technical efficiency can be accomplished with a variety of models that fall under two main rubrics: data envelop analysis (DEA) and stochastic frontier analysis (SFA). DEA methods use mathematical programming to estimate the production frontier and technical efficiency. For the most part, agricultural frontier modellers have eschewed DEA, because the method is deterministic and may not account for the diversity of shocks involved in crop production (Coelli, 1995). Moreover, DEA does not account for measurement error. For these reasons, in the following analysis we adopt a stochastic frontier approach.

Stochastic frontier analysis. SFA, developed by Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977), has been extended by numerous authors since its development.⁶ It uses distance functions to measure technical inefficiency, a notion first introduced by Shephard (1970).

Following Greene (2008), to estimate the model we assume a well-defined production structure as defined above, a single output production function, and a vector of inputs. Let

$$y \leq f(\mathbf{x}). \quad (2)$$

An output-based Debreu–Farrell measure of technical efficiency is:

$$TE(y, \mathbf{x}) = \frac{y}{f(\mathbf{x})} \leq 1. \quad (3)$$

This is also the conventional measure of total factor productivity and thus offers insight into the overall productivity of the farm.

Econometrically, in a panel setting the model can be cast as:

$$y_{it} = f(\mathbf{x}_{it}, \boldsymbol{\beta}) TE_{it} e^{v_{it}} \quad (4)$$

where $0 < TE_{it} \leq 1$, \mathbf{x} is a vector of production inputs that determine the production efficiency frontier, in our paper TE_{it} is derived from Equation (8), below. β is the vector of parameters of the production function measuring the productivity of these inputs, and v is $iid \sim N(0, \sigma_v^2)$. Using a standard Cobb–Douglas functional form, the following is estimated with maximum likelihood methods:

$$\ln(y_{it}) = \beta \ln(\mathbf{x}_{it}) + \ln(TE_{it}) + v_{it}. \quad (5)$$

A measure of technical inefficiency is thus $u_{it} = -\ln(TE_{it})$, where $u_{it} \geq 0$ and by assumption is independent of v_{it} . This is a model in which observed deviations from the production function could arise from one of two sources: productive inefficiency (u_{it}), which necessarily would be negative; or idiosyncratic stochastic effects specific to the farm (v_{it}), which can enter the model with either sign. Our specification of the error and efficiency terms follows that of Battese and Coelli's (1995) stochastic frontier model applied to panel data. The u_{it} are independently distributed as truncations at zero of the $N(m_{it}, \sigma_u^2)$. This method permits the inclusion of variables to explain differences in technical efficiency estimates. We let $m_{it} = \mathbf{z}_{it}'\delta$, where \mathbf{z}_{it} is a vector of variables that may influence the efficiency of the farm and δ is a vector of parameters to be estimated (Coelli et al., 1998). Using this method, it is possible to estimate differences in technical efficiency between large and small farms.

We estimate the following model, which maintains the Cobb–Douglas form used to test the inverse farm size relationship, keeping (the natural log of) output per hectare as the dependent variable, $\frac{Y_{it}}{T_{it}}$: β_1 :

$$\ln\left(\frac{Y_{it}}{T_{it}}\right) = \alpha_i + \beta_1 \ln(T_{it}) + \beta_2 \ln\left(\frac{L_{it}}{T_{it}}\right) + \beta_3 \ln\left(\frac{K_{it}}{T_{it}}\right) + \beta_4 \ln\left(\frac{C_{it}}{T_{it}}\right) + \beta_5 year + V_{it} - U_{it}. \quad (6)$$

Efficiency analysis is overwhelmingly empirical; the literature does not offer a convincing theory of what variables make some households more technically efficient than others. Our empirical objective is to test whether there is an inverse relationship between farm size and technical efficiency. Thus, we include farm size, T_{it} , as an explanatory variable in the inefficiency equation. We also include a vector of controls frequently incorporated into production functions, including: human capital (the education of the household head), J_{it} ; self-reported land quality and land slope (the percentage of land cultivated that is good and the percentage of land used that is flat), denoted by LG_{it} and LF_{it} , respectively; and the number of irrigated hectares cultivated, I_{it} .

In our expanded efficiency regressions we also include variables reflecting households' ability or incentives to efficiently transform inputs into output. They include transaction costs (proxied by R_{it} , a dummy variable indicating whether a village has only dirt roads); instruments for access to US migration networks (M_{it}); and ethnicity (IN_{it} , a dummy for whether the household head speaks an indigenous language). Past research finds indigenous status to be significantly and positively related to how households value traditional crops (Arslan & Taylor, 2009) and negatively related to wealth and access to resources (Perales, Benz, & Brush, 2005; Smale et al., 2003). The migration instrument is constructed through predicted values generated from a Poisson regression of the number of migrants in the household on distance to the US border, the state migration rate in 1924, the average state migration rate from 1955–1959, and weighted averages of GDP growth in migrant destination states over the five years preceding each round of the survey.⁷ Distance to the border and state migration rates have been used to instrument migration in a number of studies, on the grounds that *braceros* (the first contract workers) were loaded onto train cars in the 1940s–1950s to work in the United States, laying the foundation for future rural Mexico-to-US migration (Demirgüç-Kunt, López, Martínez, Soledad, & Woodruff, 2009; Pfeiffer et al., 2009; Woodruff & Zenteno, 2001). We include the weighted US destination growth rates on the assumption that GDP per capita in US states affects agricultural production in rural Mexican households only through its effect on migration. Labour

markets can compete with agriculture for household labour, but migrants can provide households with remittances that loosen liquidity and risk constraints on production, as well as with information. Finally, we include a dummy for the year 2007 to track changes in inefficiency over time.

The full specification of the frontier is thus:

$$\begin{aligned} \ln\left(\frac{Y_{it}}{T_{it}}\right) = & \beta_0 + \beta_1 \ln(T_{it}) + \beta_2 \ln\left(\frac{L_{it}}{T_{it}}\right) + \beta_3 \ln\left(\frac{K_{it}}{T_{it}}\right) + \beta_4 \ln\left(\frac{C_{it}}{T_{it}}\right) + \beta_5 \text{year} + \beta_6(J_{it}) + \beta_7(LB_{it}) \\ & + \beta_8(LF_{it}) + \beta_9(I_{it}) + V_{it} - U_{it}. \end{aligned} \quad (7)$$

The technical inefficiency equation is:

$$U_{it} = \delta_0 + \delta_1 \ln(T_{it}) + \delta_2(R_{it}) + \delta_3(M_{it}) + \delta_4(IN_{it}) + \delta_5 \text{year} + W_{it}. \quad (8)$$

The inverse efficiency relationship: empirical evidence. Columns 1–4 of Table 3 present the results from a Battese–Coelli SFE with time-varying inefficiency, the four inputs included in the production

Table 3. The inverse efficiency under different frontier specifications

Model	(1)	(2)	(3)	(4)
Frontier estimates	RE	RE	RE	FE
β_0	6.28*** (0.14)	6.35*** (0.14)	6.25*** (0.13)	6.25*** (0.14)
Ln (average land)	0.011 (0.038)	-0.098** (0.032)	-0.102*** (0.038)	-0.254*** (0.039)
Ln (average labour/average land)	0.194*** (0.029)	0.175*** (0.029)	0.176*** (0.027)	0.167*** (0.029)
Ln (average capital/average land)	0.046*** (0.010)	0.0326*** (0.0010)	0.0326*** (0.0098)	0.019* (0.010)
Ln (average cost/average land)	0.189*** (0.015)	0.164*** (0.015)	0.187*** (0.014)	0.059*** (0.015)
Year		0.200*** (0.066)	0.187*** (0.068)	0.090** (0.044)
Good land (% of total cultivated)		0.095 (0.075)	0.088 (0.071)	
Flat land (% of total cultivated)		0.292*** (0.077)	0.291*** (0.075)	
Irrigated land (number of hectares)		0.083*** (0.010)	0.082*** (0.010)	
Household head education			0.029*** (0.010)	
Farm fixed effects				X
Inefficiency estimates				
δ_0	-21.89*** (8.34)	-18.31** (9.09)	-17.44 (11.05)	-3.36*** (0.12)
Ln (average land)	1.01*** (0.26)	1.23 (0.76)	1.24 (1.04)	0.16** (0.064)
σ_s^2	13.68*** (4.46)	11.86** (4.74)	11.25* (5.74)	1.35*** (0.031)
γ	0.900*** (0.032)	0.897*** (0.043)	0.891*** (0.058)	0.99998*** (0.00028)
Observations	1363	1363	1361	1361
Number of households	843	843	841	841

Notes: Standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

function, and only land size in the inefficiency estimates (three random effects models and one with household fixed effects). The estimation is performed using maximum likelihood, from the software Frontier 4.1 (Coelli, 1996). The top panel of the table reports estimates of the coefficients of the equation representing the efficiency frontier, and the bottom panel presents estimated effects of household farm size on inefficiency, or distance from the frontier. U_{it} and V_{it} are expressed in terms of σ_u^2 and σ_v^2 , respectively, where the combined variance is $\sigma_s^2 = \sigma_u^2 + \sigma_v^2$ and the proportion of the total variance explained by inefficiency is $\gamma = \sigma_u^2/\sigma_s^2$. The estimate of parameter γ in all models in Table 3 is significant and close to one, indicating that the inefficiency portion of the error term is large relative to the stochastic error component, and that omitting inefficiency estimation may bias parameters in the production function.

The findings in the bottom panel of the table support the inverse efficiency hypothesis: inefficiency (the distance from the efficiency frontier) increases with farm size under both the random and fixed effects specifications. When we include the land quality variables, Model 2, the coefficient on land size is positive but not significant. This is true as well for Model 3, which includes human capital variables; it is positive and of a similar magnitude. In the fixed effects (FE) Model (4), the coefficient on land size is smaller in magnitude but significant and positive. That the parameter estimates are both positive and significant in the FE specification suggests that the omission of time-invariant variables, including land quality, is not an explanation for the inverse relationship.

These results show a clear inverse efficiency relationship. The coefficient on land size is positive in all models and highly significant in the first random effects model and more importantly in the fixed effects model; small farms appear to be producing closer to the frontier than large farms. This result occurs even after controlling for farm size in the first-stage frontier estimate. It reveals a technical efficiency relationship not captured in the inverse land productivity relationship.

All estimates on land in the first-stage frontier, except for the first random effects (RE) model, are negative and significant. However, the first-stage frontier estimates must be adjusted to reveal the true land productivity relationship. The land productivity results in the top panel represent what Battese and Broca (1997) call the elasticity of frontier output of land or the elasticity of land when production operates at its frontier. To adjust for the mean elasticity of land one must take into account the results from the inefficiency estimates on land, using the following equation:

$$\frac{\partial \ln E\left(\frac{Y_{it}}{T_{it}}\right)}{\partial \ln(T_{it})} = \beta_1 - C_{it} \left(\frac{\partial \mu_{it}}{\partial \ln(T_{it})} \right), \tag{9}$$

where μ_{it} is defined by Equation (8) and C_i is defined by

$$C_{it} = 1 - \frac{1}{\sigma} \left(\frac{\phi\left(\frac{\mu_{it}}{\sigma} - \sigma\right)}{\Phi\left(\frac{\mu_{it}}{\sigma} - \sigma\right)} - \frac{\phi\left(\frac{\mu_{it}}{\sigma}\right)}{\Phi\left(\frac{\mu_{it}}{\sigma}\right)} \right). \tag{10}$$

ϕ and Φ represent the density and distribution functions of the standard normal random variable, respectively (see Equation (9) in Battese & Broca, 1997). Table 4 reports these adjusted land productivity coefficients for the four models in Table 3.

Table 4. Inverse productivity of land

Model	(1)	(2)	(3)	(4)
Coefficient on average land	RE	RE	RE	FE
Ln (average land)	-0.174	-0.347	-0.360	-0.299

Notes: Standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Adjusted coefficients are all negative. In short, both inverse relationships are present in this model. Failure to control for efficiency effects in land size may bias tests for inverse land productivity.

Inefficiency or ignorance? Differences in estimated inefficiency across farm sizes may be misleading if unobserved variables correlate with farm size influence production. This raises a fundamental question: is inefficiency real or a reflection of researchers' ignorance? Carter (1984), and Benjamin (1995) raise similar questions in their analysis of omitted variable bias as a possible explanation for the inverse productivity relationship. The procedure to analyse inefficiency relies on decomposing the error term in the production function regression. The omission of significant variables on the frontier side inevitably influences the resulting error term and thus the inefficiency estimates. The inclusion of normally omitted inputs (for example factor quality) in the production function permits the efficiency frontier to be heterogeneous across farms. An inverse efficiency relationship might not be observed if the inclusion of a variable on the production side that is correlated with farm size removes variation in inefficiency from the error term.

The ENHRUM, unlike most data sets used to study inefficiency, provides information that can be used to test the effects of input quality and other critical omitted variables on inefficiency as well as the robustness of the inverse efficiency finding with respect to omitted variables. We added a vector of controls frequently incorporated into production functions, including: human capital (the education of the household head), J_{it} ; self-reported land quality and land slope (the percentage of land cultivated that is good and the percentage of land used that is flat), denoted by LG_{it} and LF_{it} , respectively; and the number of irrigated hectares cultivated, I_{it} . The results appear under the columns labelled 'Models 1–4' in Table 3.

Although these variables are likely to both influence crop output and be correlated with farm size, the inverse land productivity finding is robust to their inclusion in the frontier estimation: output per hectare decreases as farm size increases. The coefficients on the human capital and land quality variables have the expected signs and are strongly significant; output per hectare increases with the education of the household head, the percentage of good land, the percentage of flat land, and the number of irrigated hectares. Year effects also show a positive and significant increase in production between survey years.

The process of enriching the frontier estimates can be extended further by including farmer fixed effects. These cannot be included together with farmer human capital and land quality because the latter vary little between the two survey years, and the fixed effects control for time-invariant observables as well as unobservables. Column 4 of Table 3 (FE) reports the results, which essentially replicate the findings in the second data column of Table 2.

Logically, as these new variables are included in the production function, the errors, which are a reflection of our ignorance, decrease. The error sum of squares, a common measure of unknown variation, drops from 2,327 to 2,150 when human capital and land quality variables are included, and to 524 when we control for farmer fixed effects in the first-stage regression. Theoretically, it might be possible to reduce this result further if more were known about the determinants of production (for example time-varying unobservables); this would leave less to model as inefficiency. If the measure of the unknown were sufficiently small, SFA analysis would not be a viable approach to test for inefficiency, and indeed the meaning of inefficiency would become unclear. In the present case, there remains sufficient variation in the error, even after controlling for input quality, human capital, and farmer fixed effects, to make the farm size efficiency comparison relevant.⁸

The anatomy of inefficiency. What makes some households more technically efficient than others? Is the inverse efficiency relationship robust to the way in which inefficiency is modelled? Theory offers few guidelines on what explains inefficiency; the analysis of technical efficiency is overwhelmingly empirical. We augmented the efficiency estimation by including three additional variables; transaction costs, instruments for access to US migration networks, and ethnicity. This expanded regression is given in equations (7) and (8) and explained above.

Table 5 reports the results of this augmented inefficiency model combined with a fixed-effects estimation of the efficiency frontier. The inverse efficiency relationship holds when we control for these market and socio-demographic variables. Access to markets, including foreign labour markets, reduces inefficiency, while membership in an indigenous group does the opposite. When these three variables are included in the efficiency regression, the coefficient on land size retains its significance. We can be confident that the reason for the inverse inefficiency relationship is not due to other factors that might affect inefficiency; otherwise the result would change with the inclusion of other variables that affect efficiency. This result calls into question the conventional wisdom that large farms have an inherent efficiency advantage over smallholders.

Poor transportation infrastructure, an important correlate of transaction costs, increases inefficiency. Households with international migrants exhibit less inefficiency, consistent with migration-induced human capital and technology transfer (migrants obtain access to better agricultural management techniques and convey them to their households of origin).¹⁰ Indigenous status is positive and significant, consistent with this group's historically disadvantaged position with respect to access to information and other critical resources, as well as the high cultural value indigenous farmers place on traditional maize varieties that may not be widely marketable or highly profitable.¹¹ While the production frontier shifts out with time, farmers are, on average, more inefficient. The gamma

Table 5. Stochastic frontier estimates, full specification

	Model 5
Frontier estimates	FE
β_0	6.47*** (0.15)
Ln (average land)	-0.211*** (0.043)
Ln (average labour)	0.129*** (0.031)
Ln (average capital)	-0.007 (0.010)
Ln (average cost)	0.045*** (0.015)
Year	1.20*** (0.10)
Farm fixed effects	x
Inefficiency estimates	
δ_0	-2.75*** (0.26)
Ln (Average land)	0.257*** (0.078)
Only dirt roads	0.89*** (0.20)
International migration	-1.57*** (0.24)
Household head indigenous	0.83*** (0.21)
Year	2.38*** (0.25)
σ_s^2	0.728*** (0.024)
γ	0.99999*** (0.00022)
Observations	1348
Number of households	836

Notes: Standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 6. Likelihood ratio tests on inefficiency estimates

Model	(1)	(2)	(3)	(4)	(5)
Hypothesis	$H_0 : \gamma = \delta_0 = \delta_1 = 0$	$H_0 : \gamma = \delta_0 = \delta_1 = 0$	$H_0 : \gamma = \delta_0 = \delta_1 = 0$	$H_0 : \gamma = \delta_0 = \delta_1 = 0$	$H_0 : \gamma = \delta_0 = \delta_1 = \delta_2 = \delta_3 = \delta_4 = 0$
5% critical value ^a	7.05	7.05	7.05	7.05	11.91
Likelihood ratio statistic	10.09	14.75	14.24	789.45	377.45

Notes: ^aThe likelihood ratio statistic follows a mixed Chi-squared distribution and thus critical values are obtained from Kodde and Palm (1986), Table 1 p. 1246. Degrees of freedom are equal to the number of parameters equal to zero.

parameter for the model is practically equal to one and is highly significant, indicating that we can be confident that the estimates are correct and the variation in the error term is due almost entirely to inefficiency (Coelli, 1995). Taken together, the findings from our augmented stage-two model suggest that other than land size, access to markets, extension of household activities into non-farm sectors, indigenous status, and year effects are significant in describing efficiency in Mexican agriculture.

The elasticity of frontier output of land is negative and significant, increasing slightly with the inclusion of additional variables in the inefficiency portion of the model. The adjusted coefficient on land elasticity at mean output is -0.298 , again emphasising the inverse farm size relationship.

While we prefer the simplicity of the Cobb–Douglas model, these findings all hold up when we estimate a more flexible Translog production function, with the exception that the land coefficient, though positive, is not significant in the inefficiency portion of the model (see Online Appendix). Validation tests for the stochastic frontier regressions (Models 1–5) are shown in Table 6. Log likelihood tests of the null hypothesis that the inefficiency terms are zero indicate that all models are correctly specified and inefficiency estimates should be included in the production–frontier estimation.¹²

4. Conclusions

Recent studies question the competitiveness of small farms and call for policies to facilitate the movement of small farmers out of agriculture. Numerous studies cite an inverse relationship between farm size and land productivity; however, higher productivity does not necessarily imply greater efficiency, which is a key to being able to compete in the modern agricultural economy.

Our study confirms the existence of both an inverse productivity and an inverse efficiency relationship in rural Mexico. Fixed effects regressions and controls for input quality clearly show that small farms have a persistent advantage over large farms in terms of higher output value per hectare. Theory is generally unclear about the link between land productivity and technical efficiency. Our panel frontier analysis considers both relationships concurrently and finds evidence that the inverse relationship carries through to efficiency: other things being equal, small farms have more output value per hectare and operate closer to their efficiency frontier than large farms. These relationships are sustained even when land quality and human capital variables are included on the production side, permitting the frontier to vary across farms. They are even robust to including farm fixed effects in the production function.

The inverse efficiency relationship holds even with the inclusion of additional explanatory variables in the inefficiency component of the model. Given farm size, households' access to high quality land and education increase productivity; access to markets and US migration raise technical efficiency, while membership in indigenous groups does the opposite.

Overall, our findings offer a guardedly optimistic view of small farms' capacity to produce efficiently and potentially to adapt, despite off-farm activities' increasing demands on human resources, policy biases against small and medium farmers (Yunez-Naude, 2010), and far-reaching transformations of the agricultural supply chain. They also underline the potential importance of policies to raise technical efficiency by investing in market infrastructure, facilitating migration, and reducing economic and social-psychological barriers to efficiency on indigenous farms.

Acknowledgements

We are grateful to the William and Flora Hewlett Foundation, the National Institute of Food and Agriculture (NIFA), and the United Nations Food and Agricultural Organisation (UN-FAO) for supporting this research, and to Michael Castelhana and Fabiola Rivera for their valuable comments and assistance.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. If there were decreasing returns to scale, on the other hand, we would expect to see the widespread sub-dividing of agricultural production units over time, which usually is not the case. Moreover, as we shall see, our production function estimates using data from Mexico do not support a decreasing returns technology. Conversely, increasing returns to scale would encourage a concentration of production on ever larger farms, until productivity differences across farms vanished.
2. Carter and Yao (2002) point out that land and labour market imperfections may constrain some households but not others; this is the basis for their global test for non-separability using data from China.
3. Formally this is known as input-oriented efficiency, defined as the maximum equi-proportionate reduction in all inputs that is feasible with given technology and outputs. Output-oriented efficiency is defined as the maximum radial expansion in all outputs that is feasible with given technology and inputs. These two measures are the same for a constant-returns-to-scale technology (Debreu, 1951; Farrell, 1957).
4. Even eliminating the tails, 5 per cent of the bottom and top of the distribution of operated land size leads to a significant downward trend and inverse relationship (additional graph and regression available upon request).
5. Capital is calculated as reported animal or machinery rental cost if rented in or if own use, imputed at the median per hectare cost from village level medians (if insufficient observations occur at the village level, we impute at increasing geographic levels). All capital costs and purchased inputs are increased by MXN 1 to avoid loss of observations when logged.
6. See Fried, Lovell, and Schmidt (2008) for a comprehensive discussion of current methods.
7. We use the shares of migrants from each household (i) in each state (j) in the survey year as weights, where j are indices for states in the United States. That is,

$$migIV_{GDP} = \sum_{(j=1)}^{51} Share Migs_{ij} \times GDPgrowth_j^{preceding 5 years}.$$

See also Arslan and Taylor (2012) for more details on this instrument.

8. This is also confirmed later in Table 6, which reports that log likelihood tests of only the production function regressions are rejected in favour of the full inefficiency regressions for every model specification.
9. We use the shares of migrants from each household (i) in each state (j) in the survey year as weights, where j are indices for states in the United States. That is,

$$migIV_{GDP} = \sum_{(j=1)}^{51} Share Migs_{ij} \times GDPgrowth_j^{preceding 5 years}.$$

See also Arslan and Taylor (2012) for more details on this instrument.

10. It is also possible that migration-induced labour shortages create incentives to reorganise their production activities around labour scarcity and produce closer to the frontier.

11. Arslan and Taylor (2009) find that, in indigenous regions of Mexico in particular, farmers excessively value traditional maize varieties and persist in producing them despite falling prices in the market.
12. In Models 1–4 with three restrictions the likelihood ratio statistic is greater than the critical value of 7.05. Likewise, Model 5 with four restrictions has a test statistic of 320.52, which is significantly greater than the critical value of 11.91.

References

- Aigner, D., Lovell, C., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, 21–37. doi:10.1016/0304-4076(77)90052-5
- Akerloff, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115, 715–753. doi:10.1162/qjec.2000.115.issue-3
- Arslan, A., & Taylor, J. E. (2009). Farmers' subjective valuation of subsistence crops: The case of traditional maize in Mexico. *American Journal of Agricultural Economics*, 91, 956–972. doi:10.1111/j.1467-8276.2009.01323.x
- Arslan, A., & Taylor, J. E. (2012). Transforming rural economies: Migration, income generation and inequality in rural Mexico. *Journal of Development Studies*, 48, 1156–1176. doi:10.1080/00220388.2012.682985
- Baily, M., & Gersbach, H. (1995). Efficiency in manufacturing and the need for global competition. *Brookings Paper on Economic Activity: Microeconomics*, 1995, 307–347.
- Barrett, C. (1996). On price risk and the inverse farm size-productivity relationship. *Journal of Development Economics*, 51, 193–215. doi:10.1016/S0304-3878(96)00412-9
- Barrett, C. B., Reardon, T., & Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: Concepts, dynamics, and policy implications. *Food Policy*, 26, 315–331. doi:10.1016/S0306-9192(01)00014-8
- Battese, G. E., & Broca, S. S. (1997). Functional forms of stochastic frontier production functions and models for technical inefficiency effects: A comparative study for wheat farmers in Pakistan. *Journal of Productivity Analysis*, 8, 395–414. doi:10.1023/A:1007736025686
- Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20, 325–332. doi:10.1007/BF01205442
- Benjamin, D. (1995). Can unobserved land quality explain the inverse productivity relationship? *Journal of Development Economics*, 46, 51–84. doi:10.1016/0304-3878(94)00048-H
- Bhalla, S. S., & Roy, P. (1988). Mis-specification in farm productivity analysis: The role of land quality. *Oxford Economic Papers*, 40, 55–73.
- Boucher, S. R., Stark, O., & Taylor, J. E. (2009). A gain with a drain? Evidence from rural Mexico on the new economics of the brain drain. In J. Kornai, L. Matyas, & G. Roland (Eds.), *Corruption, development and institutional design* (pp. 100–119). Basingstoke: Palgrave.
- Carter, M. (1984). Identification of the inverse relationship between farm size and productivity: An empirical analysis of peasant agricultural production. *Oxford Economic Papers*, 36, 131–146.
- Carter, M., & Yao, Y. (2003). Local versus global separability in agricultural household models: The factor price equalization effect of land transfer rights. *American Journal of Agricultural Economics*, 84, 702–715. doi:10.1111/1467-8276.00329
- Chayanov, A. V. ([1926] 1991). *The theory of peasant co-operatives* (D.W. Benn & V. Danilov Trans.). Columbus: Ohio State University Press.
- Coelli, T. (1995). Recent developments in frontier modelling and efficiency measurement. *Australian Journal of Agricultural Economics*, 39, 219–245. doi:10.1111/j.1467-8489.1995.tb00552.x
- Coelli, T. (1996). A guide to Frontier Version 4.1.: A computer program for stochastic frontier production and cost function estimation (Working Paper 96/07). Brisbane: Centre for Efficiency and Productivity Analysis.
- Coelli, T., Rao, D. Prasada, & Battese, G. (1998). *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic Publishers.
- Collier, P. (2008). The politics of hunger: How illusion and greed fan the food crisis. *Foreign Affairs*, 87, 67–79.
- Debreu, G. (1951). The coefficient of resource utilization. *Econometrica*, 19, 273–292.
- Demingüç-Kunt, A., López, E., Martínez, M. S., Soledad, M., & Woodruff, C. (2009). Remittances and banking sector breath and depth: Evidence from Mexico, policy research (Working Paper Series no. 4983). Washington, DC: The World Bank.
- Docquier, F., Faye, O., & Pestieau, P. (2008). Is migration a good substitute for education subsidies? *Journal of Development Economics*, 86, 263–276. doi:10.1016/j.jdeveco.2007.12.007
- Docquier, F., & Marfouk, A. (2006). International migration by education attainment, 1990–2000. In Ç. Özden & M. Schiff (Eds.), *International migration, remittances, and the brain drain*. Washington, DC: World Bank.
- Eswaran, M., & Kotwal, A. (1986). Access to capital and agrarian production organization. *Economic Journal*, 96, 482–498.
- Fang, H., & Lounsbury, G. C. (2005). “Dysfunctional identities” can be rational. *The American Economic Review*, 95, 104–111. doi:10.1257/000282805774669998
- Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A, General*, 120, 253–282.
- Fried, H. O., Lovell, C. A., & Schmidt, S. S. (2008). Efficiency and Productivity. In H. O. Fried, C. A. Lovell, & S. S. Schmidt (Eds.), *The measurement of productive efficiency and productivity change* (pp. 3–91). New York, NY: Oxford University Press.

- Green, A., & Mayes, D. (1991). Technical inefficiency in manufacturing industries. *Economic Journal*, 101, 523–538.
- Greene, W. H. (2008). The econometric approach to efficiency analysis. In H. O. Fried, C. A. Lovell, & S. S. Schmidt (Eds.), *The measurement of productive efficiency and productivity growth* (pp. 92–250). New York, NY: Oxford University Press.
- Griffiths, R. (2001). Product market competition, efficiency and agency costs: An empirical analysis (Working Papers W01/12). London: Institute for Fiscal Studies.
- Heltberg, R. (1998). Rural market imperfections and the farm-size productivity relationship: Evidence from Pakistan. *World Development*, 26, 1807–1826. doi:10.1016/S0305-750X(98)00084-9
- Kodde, D. A., & Palm, F. C. (1986). Wald Criteria for jointly testing equality and inequality restrictions. *Econometrica*, 54, 1243–1248.
- Meeusen, W., & van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18, 434–444.
- Oxoby, R. J. (2004). Cognitive dissonance, status and growth of the underclass. *The Economic Journal*, 114, 727–749. doi:10.1111/eoj.2004.114.issue-498
- Perales, H. R., Benz, B. F., & Brush, S. B. (2005). Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. *Proceedings of the National Academy of Sciences*, 102, 949–954. doi:10.1073/pnas.0408701102
- Pfeiffer, L., Lopez-Feldman, A., & Taylor, J. E. (2009). Is off-farm income reforming the farm? Evidence from Mexico. *Agricultural Economics*, 42, 125–138. doi:10.1111/j.1574-0862.2009.00365.x
- Pingali, P., Khwaja, Y., & Meijer, M. (2005). Commercializing small farms: Reducing transaction costs (Working Paper No. 05-08). Rome: FAO-ESA.
- Sen, A. (1966). Peasants and dualism with or without sur-plus labor. *The Journal of Political Economics*, 74, 425–450.
- Shephard, R. W. (1970). *Theory of Cost and Production Functions*. Princeton: Princeton University Press.
- Smale, M., Bellon, M. R., Aguirre Gomez, J. A., Manuel Rosas, I., Mendoza, J., Solano, A. M. . . . Berthaud, J. (2003). The economic costs and benefits of a participatory project to conserve maize landraces on farms in Oaxaca, Mexico. *Agricultural Economics*, 29, 265–275. doi:10.1111/agec.2003.29.issue-3
- Woodruff, C., & Zenteno, R. (2001). Remittances and microenterprises in Mexico (Working Paper). San Diego, CA: University of California at San Diego.
- Yunez Naude, A. (2010). Las políticas públicas al sector rural: El carácter de las reformas para el cambio estructural. In Antonio Yúnez Naude Coordinator, *La Economía Rural*, Vol. XI of *Los grandes problemas de México*. In M. Ordorica y J. F. Prud'homme (Eds.) Mexico City: El Colegio de Mexico.