Property Rights and Efficiency:
Evidence from Mexico's Land Titling Program

Travis McArthur*

Abstract

Secure property rights over land facilitates entry into and exit from the agricultural sector. In an environment with heterogeneity in agents' skill levels in agricultural and non-agricultural activities, enabling land transactions will shift the skill composition of the agricultural sector, altering the efficiency of agricultural production. Establishing secure property rights generates ambiguous outcomes for agricultural efficiency. Once land titles are granted, the best farmers may choose to leave farming to capture higher incomes in another occupation. To explore this issue empirically I examine Mexico's massive land titling program implemented nationwide between 1993 and 2006. I do not find any change in agricultural efficiency caused by this program. However, the analysis reveals strong evidence for a rise in the number of individuals exiting and entering farming.

*Department of Agricultural & Applied Economics, University of Wisconsin-Madison. Email: tdmcarthur@wisc.edu.
Previously circulated under the title "Direct Measurement of Efficiency Gains from Land Titling: PROCEDE's Effect upon the Productivity of Mexican Agriculture." This paper benefitted from input from Jennifer Alix-Garcia, Bradford Barham, Jean-Paul Chavas, Alain de Janvry, Gustavo Gordillo, Cornelia Ilin, Alexey Kalinin, Matthew Klein, Daniel Phaneuf, Rashesh Shrestha, Christopher Taber, Emilia Tjernström, Srinivasan Vasudevan, and Natalia Volkow. Data analysis of the Mexican agricultural censuses occurred on the premises of INEGI, the national statistics agency of Mexico, to protect respondent confidentiality. This version December 2016.

1 Introduction

From 1993 to 2006, the government of Mexico rolled out a massive program to grant individual land titles to over 3.6 million farmers. I use this natural experiment to investigate longstanding questions about the relationship between secure property rights and economic efficiency. The main contribution of this paper is providing a measurement of the effect of the Mexican land titling program on efficiency in the agricultural sector. Another contribution is development of a two-sector model with heterogenous agents and a tradeable sector-specific factor of production.

Cross-country evidence suggests that secure property rights is an important determinant of economic growth and efficiency (de Soto 2000). Rigorous identification of the impact of property rights is a challenging research proposition, however. Property rights institutions are typically nationwide, so associations between growth and property rights may be confounded by country-level effects. To analyze the causal relationship in my particular context I construct a theoretical model that yields ambiguous predictions about the relationship between a land titling program and efficiency. I then exploit the staggered rollout of the Mexican program, known as PROCEDE, to estimate its effect on efficiency and the volume of land transactions.

The theoretical framework developed in this paper assumes that the land titles permit economic agents to pursue their comparative advantage. When land rights become transferable, agents can choose to switch between the two economic sectors: agriculture and off-farm activities. Define farming skill as technical efficiency in production, meaning the ability to extract more output from a farm once the production technology and input levels have already been accounted for. In an environment where agents have heterogeneous skill levels, any sector-switching by agents has consequences for efficiency. To analyze these consequences, I extend the Roy (1951) model of occupational choice to the case where agents acquire property rights for sector-specific factors of production.

What predictions does this model generate? After secure property rights are established, agents in the model do what is best for themselves. They do not necessarily do what is best for the efficiency of the sector that they previously inhabited. Using the framework of my model, I prove that a rise or decline in efficiency of the agricultural sector are both possible outcomes when the government establishes secure property rights.

To obtain some intuition on how such a result can come to pass, imagine that an agent is better than average at both farming and off-farm activities – an absolute advantage in both sectors – but she finds that she can make more money in the city after selling her land. When she obtains title to her land, liquidates it, and exits farming, the mean skill level in the agricultural sector falls because the remaining farmers are not as skilled as her, on average. If she had had an absolute disadvantage in farming, on the other hand, her decision to abandon farming would have raised agriculture's mean skill level. Despite the implications of the model for efficiency in one particular sector, economy-wide efficiency rises when secure land rights are granted. Thus the model still fulfills the expectations of standard economic reasoning.

Next I empirically investigate the impact of the Mexican land titling program on agricultural efficiency. I achieve identification of the treatment effect by taking advantage of the rollout of the program over time and space. Using agricultural census data, I construct a panel of agricultural efficiency and estimate a fixed-effects model. This approach yields a consistent estimator of the treatment effect under the assumption that there is no unobserved factor that varies across both time and space and is correlated with both the program rollout and efficiency.

The estimated treatment effect is not statistically different from zero. Notwithstanding the null results regarding efficiency, I find strong evidence that the program stimulated transactions activity in the land market. Administrative records of a government subsidy program reveal a surge in the volume of land transfers and a rise in the number of individuals exiting and entering farming in the three years following land titling. Since the theoretical model predicted sector-switching by agents but an ambiguous outcome for efficiency, the results of the estimation are consistent with theory.

A large body of work has analyzed the impact of land titling on the incentive to make land more productive by investing in it (Galiani & Schargrodsky 2010; Besley 1995; Feder 1988). My work speaks to the less-explored question of whether land becomes more productive merely by virtue of changing hands, controlling for any investment changes. Prior work on land titling and land market liberalization has found substantial gains in agricultural efficiency, in contrast to my finding (Holden, Deininger, & Ghebru 2009; Jin & Deininger 2009). I claim that my approach permits better identification of the causal effect than previous studies. Existing studies of the Mexican program have examined its impact upon migration, land transactions, and credit access, but no work has directly addressed the efficiency questions that I explore in this paper (de Janvry et al. 2015; World Bank 2001; Johnson 2001).

The paper first discusses the institutional context of the land titling program in Mexico. Then it reviews prior work on land titling and technical efficiency issues. Next comes the theoretical model. Then I outline a data generating process that describes how inputs and differing efficiency levels combine to produce agricultural output. An estimating equation for this data generating process, which aims to measure the treatment effect of land titling on efficiency, follows. Finally, I examine evidence of the program's effect on land transactions, land concentration, and receipt of credit.

2 Background

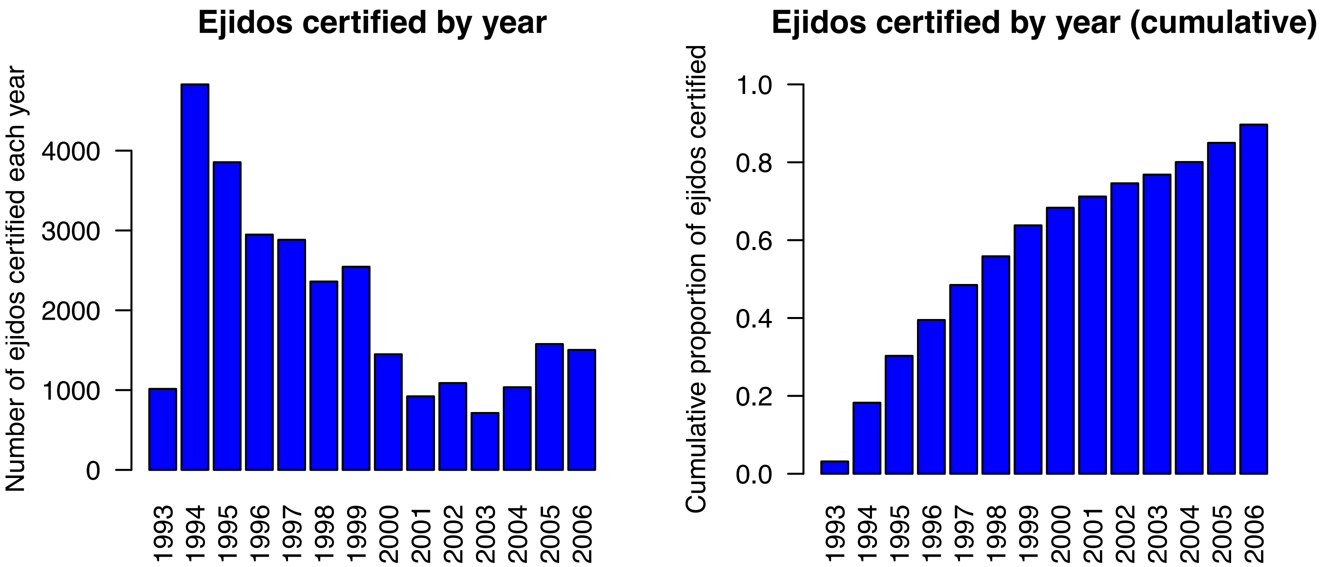
The Mexican Revolution, fought partially over inequality in landholdings, completely disrupted Mexican society and led to the deaths of 10 percent of the population (McCaa 2001). One of main outcomes of the Revolution was the creation of the legal framework in 1917 for communities called ejidos. The purpose of the ejido system was to redistribute land from large landholders to landless individuals. A secondary goal was to maintain a clientelistic relationship between the one-party Mexican state and rural people. To establish an ejido, a group of individuals would create a petition for land. The state governor and the Mexican president would then grant a piece of land to this group, expropriating private land if deemed necessary. New ejidos have been created throughout the post-Revolution period. Eventually, over half of all arable land in Mexico became part of one of the 30,000 ejidos created in this way (World Bank 2001). Ejidos are the unit of observation in this paper.

Ejidos consist of three types of landholdings. "Land for common use" includes livestock grazing land, forests, and bodies of water. "Urban plots" are for dwellings. Finally, usufruct rights to agricultural parcels for individual use were issued. These individual parcels are my object of study. In theory, members of the ejido were supposed to have been granted certificates describing their parcel usufruct rights, but a survey in 1983 found that 86 percent of ejido members lacked these certificates (Heath 1992). About 3 percent of ejidos were completely collectivized (Heath 1992). In my analysis I will ignore these ejidos because their issues are distinct from the typical ejido.

In keeping with the goal of avoiding land re-consolidation, the laws governing ejidos severely restricted transfer of the parcel usufruct rights. Sale and rent of land was prohibited, and a two-year absence from the ejido was grounds for expropriation and re-allocation of an individual's land (World Bank 2001). Land could be transferred via inheritance, but to prevent land fragmentation only one heir could inherit the land (Nuijten 2003). Hiring labor to work the land was not allowed. Although land transactions were illegal, there was an active black market in ejido land prior to the 1990's reforms (Procuraduría Agraria 1998). To the extent that this black market was widespread, the efficiency effect that I intend to measure may be attenuated.

In the midst of a series of free-market reforms that included NAFTA, the Mexican legislature altered ejido rules to permit the sale of parcels to other members of a particular ejido and the rent of parcels to anyone. The prohibition on hiring labor to work ejido land was also lifted (de Ita 2006). The amendments were approved in late 1991 and became law in January 1992 (Johnson 2001). Part of the motivation behind the land market liberalization was to allow transfer of land to those who are better at farming and thereby

Figure 1



Source: Padrón e Historial de Núcleos Agrarios (PHINA) Database

boost efficiency (de Ita 2006). Typically, identification of the effect of a nationwide institutional change is challenging because it is difficult to disentangle the effect of the policy shift from other factors that are changing over time. However, in this case the change in the law was effectively operationalized by a program whose implementation varied across time and space: PROCEDE.

The program (Programa Nacional de Certificación de Derechos Ejidales y Solares Urbanos — National Certification Program of Ejido Rights and Urban Lots) granted land titles to members of ejidos. Technically, ejido members could sell or rent land without the titles after the 1992 reform, but the land titles increased the legal security of the transaction. Teams from the central Mexican government made contact with ejidos and explained the purpose and operational details of PROCEDE. Once the ejido leadership gave initial approval, government agents measured parcels with GPS technology and drew up certificates for each ejido member. A final ejido assembly vote approved the parcel delimitations, and certificates were issued to all members of a given ejido simultaneously. The simultaneous, compulsory nature of the land titling removes concerns about individual endogenous selection into the program. By the end of the program in 2006, 91 percent of all ejidos had been given land titles. Figure 1 shows the progression of PROCEDE rollout over time.

In practice, the timing of the implementation of the program in each ejido was the result of the interplay between central government agents and ejido governance structures. Previous work has shown that timing was associated with observable characteristics such as ejido size (de Janvry et al. 2015). With my agricultural production data I find that ejidos with earlier implementation dates had, at the 1991 baseline, a lower number of farmers per ejido, higher output per farm, larger number of hectares per farm, and higher crop yield. A fixed-effects strategy will remove any endogenous selection bias that results from time-invariant observable or unobservable characteristics of each ejido. If the timing of implementation of the program in certain ejidos was correlated with long-term upward or downward trends in productivity, the fixed effects would not fully eliminate endogeneity bias. The fixed effects handles only levels, not trends.

3 Prior work

Academic study of the productivity impacts of PROCEDE has focused on the indirect correlates of productivity growth. A World Bank report on the ejido reforms did not detect any shifts in land sales due to PROCEDE (World Bank 2001). On the other hand, the report found that it caused more activity in the land rental market, although the lease terms were almost all less than one year. Johnson (2001) did not find any effect of PROCEDE on access to credit, which could have boosted investment.

The most important study in this area is de Janvry et al. (2015). It finds that PROCEDE accounted for 20 percent of the outmigration from ejido communities. The paper also found that PROCEDE led to a lower number of farmers in particular, although the estimated effect is not very precise. The total land under cultivation in ejidos seemed to be unaffected.

If even a small proportion of the PROCEDE-related migrants left due to having a comparative disadvantage in farming, the change in mean technical efficiency in early PROCEDE ejidos should be detectable. My identification strategy is similar to de Janvry et al. (2015) in that their method relies upon fixed effects and the rollout of the program over time and space. Their identifying assumption is that any ejido-specific characteristics that vary over time are uncorrelated with at least one of 1) the timing of PROCEDE implementation; or 2) migration away from ejidos. My identifying assumption is that any ejido-specific characteristics that vary over time are uncorrelated with at least one of 1) the timing of PROCEDE implementation; or 2) changes in technical efficiency.

Work on the technical efficiency impacts of land tiling and land transactions includes Holden, Deininger, & Ghebru (2009) and Jin & Deininger (2009). The former study finds that a land titling program in Ethiopia — where rent but not sale of land was permitted — raised technical efficiency by 45 percent. They use a propensity score matching technique to handle the endogeneity of selection into land titling. The latter study estimates the increase in efficiency associated with land transactions in China. Via estimation of a panel data model, they find an average rise in efficiency of 60 percent, but they make no causal claim about the impacts of rental markets themselves.

This analysis of the efficiency effects of PROCEDE fits into a broader research agenda. In his survey of the state of productivity research, Syverson (2011) identifies measurement of the productivity impact of policies in developing countries as a research frontier:

> While research has identified misallocation as a source of the problem, it hasn't really pinned down exactly what distortions create gaps between the social marginal benefits and costs of inputs across production units. It is hard to implement policies that close these gaps and the variation between them (i.e., reallocate inputs more efficiently) without knowing the nature of the gaps in the first place. That said, there has been some early progress on this front. Witness the efforts to tie misallocation to various labor market policies. Much remains to be done, however, and this is an important area for further effort.

Productivity differences across firms can be truly huge. Hsieh & Klenow (2009) find Indian and Chinese manufacturers in the 90th percentile of efficiency are four times as productive as those in the 10th percentile. In other words, those firms at the top of the distribution are producing four times as much output with the same inputs as those at the bottom. Policies can greatly influence the average efficiency level of firms. Kalirajan, Obwona, & Zhao (1996) find that Chinese agricultural total factor productivity rose by an annual average of 8 percent for six years (1978-1984) during the decollectivization of the Chinese agricultural sector.

Finally, this research can find a place in the "structural transformation" literature. Lewis (1954) was an early proponent of this approach to development macroeconomics. The idea is to encourage "surplus" labor to leave the underperforming agricultural sector and join the "modern" manufacturing and services sector. Liberalization of the land market could accelerate this transition.

4 Theoretical model

I first outline the intuition of the theoretical model. Assume that farmers are heterogeneous in their skill levels. Farmers with a higher skill level are more technically efficient. In fact let skill level be synonymous with technical efficiency. Technically efficient farmers are closer to the production possibility frontier. They produce more output than their peers when choosing the same level of inputs as their peers. There is some off-farm income option, which includes activities within the farming community as well as migration to urban areas or the United States. Say that the farmers have heterogenous outcomes in this off-farm option. Hence, some farmers have a comparative advantage in farming and others have it in off-farm activities. For simplicity, I will assume that individuals either farm exclusively or generate income exclusively through off-farm activities; there is no diversification. I will further assume that only individuals exist, not households, so there are no household-level risk diversification strategies that could generate a complex portfolio of occupations.

Without land titles, land transactions would be limited in such an environment. Landowners who wish to concentrate their efforts on off-farm activities would not see any rental or sale income from shedding their land. Furthermore, those with a comparative advantage in farming would not be so keen to collect land abandoned by the first group because defending a claim to the newly-acquired land would, in turn, be challenging. Now grant land titles to all those with usufruct land rights. With land titles the incentives to transact land would shift so as to put more land in the hands of those with a comparative advantage in farming.

For the moment assume that farming skill level and off-farm skill level are negatively correlated across farmers. In this case, one set of ejido members would have an absolute advantage in farming and another set an absolute advantage in off-farm activities. As more land shifts to ejido members who, in an absolute sense, are more technically efficient in farming, the average agricultural technical efficiency of the ejido will rise.

Alternatively, assume that farming skill level and off-farm skill level are positively correlated across farmers. Therefore, there is one set of farmers that is better at both farming and off-farm activities than all the other farmers. With titling, land would still change hands because some farmers would have a comparative advantage in farming. Who ultimately owns the land would be an open question. If the high-skill people accumulate more land, technical efficiency should rise. If low-skill people accumulate more land, technical efficiency could fall.

Technical inefficiency as an additive term in the profit function

I now turn to a mathematical presentation of the model. This section will discuss how technical inefficiency can fit in a Roy model selection framework. It will then show that, due to Roy-type selection phenomena, the effect of land titling upon average agricultural technical efficiency may be positive or negative.

Let the production relation be

$$y_j = f(\mathbf{x_j}) \cdot \exp\{u_j\}$$

where

y_j is output for the j'th farm,

$f(\cdot)$ is the production function common to all firms.

$\mathbf{x_j}$ is the vector of inputs employed by the j'th firm, and

u_j is the value of technical inefficiency of the j'th firm. We must have $u_j \leq 0$. The j'th firm is fully efficient, i.e. operating at the technological frontier, if $u_j = 0$.

Kumbhakar (2001) shows that, under the assumption that $f(\cdot)$ is homogeneous of degree r and that $r < 1$, the profit function can be represented as the product of u_j and some terms that do not depend on u_j, i.e. the profit function is multiplicatively separable in u_j. This will be a crucial representation for the goal of incorporating technical inefficiency into the standard form of the Roy model. This representation is

$$\ln(\pi^a) = \frac{1}{1-r}\ln(p) + \ln(G(\mathbf{w})) + \frac{1}{1-r} \cdot u_j \tag{1}$$

where

π^a is actual (observed) profit,

p is the output price,

$\mathbf{w}$ is a vector of input prices, and

$G(\cdot)$ is some function that is homogeneous of degree $-\dfrac{r}{1-r}$ in $\mathbf{w}$

It is easy to see from a simple algebraic standpoint that we need $r < 1$ because $r = 1$ results in division by zero and if $r > 1$ there is, on the one hand, higher output prices leading to lower profit and, on the other hand, greater magnitudes of u_j (recall $u_j \leq 0$) resulting in greater profit. Decreasing returns to scale are probably satisfied in this context of small-scale farming in Mexico. Even if the underlying technology exhibits constant returns to scale, the inability to contract on labor effort or monitor labor — the principal-agent problem — ensures that hired labor is "lower quality" than family labor. Hence, once family labor is exhausted, decreasing returns to scale set in.

The Roy model of occupation selection

As a foundation, I will use the basic Roy model as described in French & Taber (2011) and Borjas (1989). Then I will build upon the basic model to incorporate firm-owning agents who wish to maximize profits and, finally, comparative statics dealing with land titling.

Let us start with earnings in two activities: building houses B and farming F. Housebuilding is a stand-in for any off-farm labor activity in Mexico. Assume that individuals are risk-neutral or assume that there is no uncertainty in returns over time. Individuals are free to do either activity or both activities, but we will see that the two activities are effectively mutually exclusive due to the incentives that the individuals face. Let Y_{Bj} be the j'th individual's potential log earnings from building and let Y_{Fj} be the j'th individual's potential log earnings from farming. These are potential earnings rather than realized earnings since ultimately each individual will choose only one of the two activities. The Roy model is:

$$\begin{aligned} Y_{Bj} &= \mu_B + \epsilon_{Bj} \\ Y_{Fj} &= \mu_F + \epsilon_{Fj} \end{aligned} \tag{2}$$

where

Y_{Fj} and Y_{Bj} are possible log earnings from building and farming for individual j

μ_B and μ_F are mean earnings of building and farming, and

ϵ_{Bj} and ϵ_{Fj} are random variables whose realizations vary across individuals j. These represent inherent skill or ability in the two activities.

ϵ_{Bj} and ϵ_{Fj} are distributed jointly normal:

$$\begin{bmatrix} \epsilon_{Bj} \\ \epsilon_{Fj} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_B^2 & \sigma_{BF} \\ \sigma_{BF} & \sigma_F^2 \end{bmatrix} \right) \tag{3}$$

Under the assumptions that individuals 1) only seek to maximize earnings, 2) know μ_B and μ_F, and 3) know their realized ϵ_{Bj} and ϵ_{Fj}, the actual occupational choice C_j of the j'th individual is simply

$$C_j = \begin{cases} B & \text{if } Y_{Bj} > Y_{Fj} \\ F & \text{if } Y_{Bj} \leq Y_{Fj} \end{cases}$$

The key question that this framework will help us answer is whether the individuals with the highest realization of ϵ_{Fj} are more likely to farm. In other words, do the best farmers actually farm and do the best builders build? Note that we must have $\sigma_F^2, \sigma_B^2 > 0$ since there is no "best" or "worst" group if the agents all have an equal skill level in a particular activity. And let $\sigma_{BF} \neq 0$ because there is nothing interesting to say if the two skills are completely unrelated.

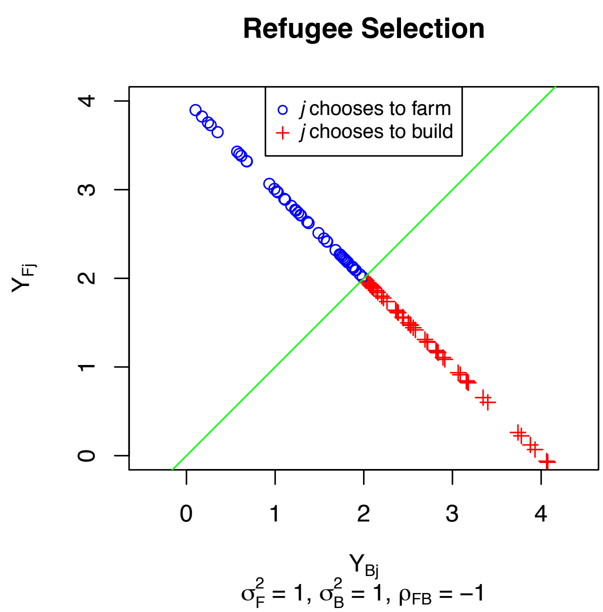
There are three possible scenarios depending on how agents are selected into farming. The first one, called refugee selection by the migration literature, occurs when the best farmers farm and the best builders build. The second, called positive selection into farming, occurs when the best farmers farm, but the best farmers are also the best builders, so the agents that are not so great at building end up choosing to build. The third option is negative selection into farming in which the best farmers are even better at building, so they choose to build; the not-so-great farmers choose to farm. If negative selection into farming prevails, there is positive selection into building, and vice versa. There are no general equilibrium interactions between agents here.

Define $\bar{\rho} \equiv \min\left\{\frac{\sigma_B}{\sigma_F}, \frac{\sigma_F}{\sigma_B}\right\}$. The formal conditions for these selection outcomes are

$$
\begin{aligned}
\rho_{BF} < \bar{\rho} \quad &\Longrightarrow \quad \text{refugee selection} \\
\rho_{BF} > \bar{\rho} \ \wedge \ \sigma_F > \sigma_B \quad &\Longrightarrow \quad \text{positive selection} \\
\rho_{BF} > \bar{\rho} \ \wedge \ \sigma_F < \sigma_B \quad &\Longrightarrow \quad \text{negative selection}
\end{aligned}
$$

Hence, the type of selection depends on all of the values of the elements of the variance-covariance matrix of ϵ_{Bj} and ϵ_{Fj}. Stated simply, these conditions require that the correlation between ϵ_{Bj} and ϵ_{Fj} be negative or sufficiently low for refugee selection to occur. If refugee selection can be ruled out, positive selection will prevail if the variance of farming skill is higher than that of building skill. Negative selection will occur if variance of farming skill is lower than that of building skill.

Figure 2



Refugee selection probably does not arise often in reality since it would require ability in one activity to translate poorly into ability in the other. It can often require that the two types of abilities are nearly inversely related. Figure 2 illustrates this refugee selection.[1] I have chosen $\rho_{BF} = -1$ for simplicity. The plot is in Y_{Bj}, Y_{Fj} space for the realized

[1] A lecture by Christopher Taber inspired Figures 2–6.

values of ϵ_{Bj} and ϵ_{Fj}. The plot symbol indicates what occupation each individual will choose. A blue circle indicates that the agent's income from farming would be higher than that from building. A red cross indicates the opposite. In green is the 45-degree line, which represents the cutoff for the agents' decision rule. Figure 3 shows positive selection, while Figure 4 shows negative selection. When $\rho_{BF} = 1$, positive selection into farming will occur if $\sigma_F^2 > \sigma_B^2$ and negative selection will occur if $\sigma_F^2 < \sigma_B^2$.

The effect of land titling

Now that I have stated the setup of the traditional Roy model and its main theoretical implications, I will explore the effects of land titling. This will involve starting at an equilibrium where all agents are not free to choose occupations. After PROCEDE is implemented, an equilibrium with a liberalized land market is established in which agents select into their preferred occupation.

Using (1), re-express Y_{Fj} as farm profit

$$Y_{Fj} = \ln\left(\pi^a\right) = \frac{1}{1-r}\ln\left(p\right) + \ln\left(G\left(\mathbf{w}\right)\right) + \frac{1}{1-r}\cdot u_j$$

We can immediately return to the familiar environment of (2) by re-centering and re-scaling via these definitions:

$$\begin{aligned} \mu_F &\equiv \frac{1}{1-r}\ln\left(p\right) + \ln\left(G\left(\mathbf{w}\right)\right) \\ \epsilon_{Fj} &\equiv \frac{1}{1-r}\cdot u_j \end{aligned} \tag{4}$$

This produces $Y_{Fj} = \mu_F + \epsilon_{Fj}$. Farms are price takers so their input price vector $\mathbf{w}$ and output price p are common to all farmers within a given ejido.

Now consider the situation in ejidos in 1991 before PROCEDE. For simplicity, assume that land exists as indivisible units and that one unit is necessary and sufficient for farm production. Denote variables associated with these individuals with superscripts R for "rights". There are other individuals in the ejido or nearby environs who do not have these rights. Use the N superscript for "no rights". Let the random vectors $\left[\begin{array}{cc}\epsilon_{Bj}^R & \epsilon_{Fj}^R\end{array}\right]$ and $\left[\begin{array}{cc}\epsilon_{Bj}^N & \epsilon_{Fj}^N\end{array}\right]$ be independently and identically distributed as in (3).

Denote occupational earnings and choices in this pre-PROCEDE equilibrium with " $^\circ$ ".

Given that the farmers with usufruct rights have a choice to farm or not farm, we know that $\mathring{Y}_{Bj}^R \leq \mathring{Y}_{Fj}^R$ if and only if $\mathring{C}_j^R = F$.

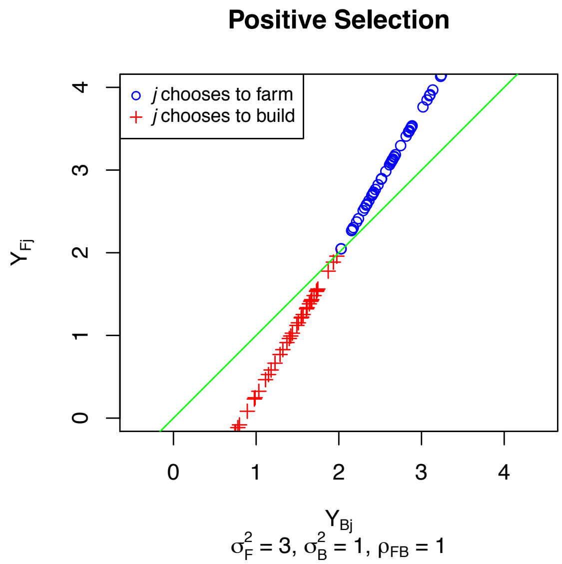
We do not know the ordering of $\mathring{Y}^N_{Bj}$ and $\mathring{Y}^N_{Fj}$ for each individual with no rights because they have not been given the opportunity to farm.

Now give land titles to those who had usufruct rights. These agents can now liquidate their landholdings. Denote occupational earnings and choices in this post-PROCEDE equilibrium state with "*". Their occupation decision $\overset{*}{C}{}^R_j$ is then determined by

$$\begin{aligned} \overset{*}{Y}{}^R_{Bj} &= \mu_B + L + \epsilon^R_{Bj} \\ \overset{*}{Y}{}^R_{Fj} &= \mu_F + \epsilon^R_{Fj} \end{aligned} \quad , \quad \overset{*}{C}{}^R_j = \begin{cases} B & \text{if } \overset{*}{Y}{}^R_{Bj} > \overset{*}{Y}{}^R_{Fj} \\ F & \text{if } \overset{*}{Y}{}^R_{Bj} \leq \overset{*}{Y}{}^R_{Fj} \end{cases}$$

where L is the market value of their land unit. Since Y is log earnings, L enters multiplicatively rather than linearly into the equation for potential earnings from building. L may enter multiplicatively if credit markets work imperfectly and a larger initial capital stock scales up income from the off-farm activity. Another way to map the model onto reality is to have Y be level, rather than log, earnings, so L would be just a non-productive income boost.

Figure 3



The earnings from switching to the building occupation have risen while earnings from farming have not changed. To ensure that there exist farmers sufficiently close to the knife-edge decision rule, assume there are an infinite number of these agents with unit mass. Given that $L > 0$, some of these agents will switch to $\overset{*}{C}{}^R_j = B$. This set of agents will sell off their land and become builders.

The "no rights" group, on the other hand, must buy or rent land in order to engage in farming, so L is subtracted from Y^N_{Fj} in the post-PROCEDE decision rule:

Figure 4

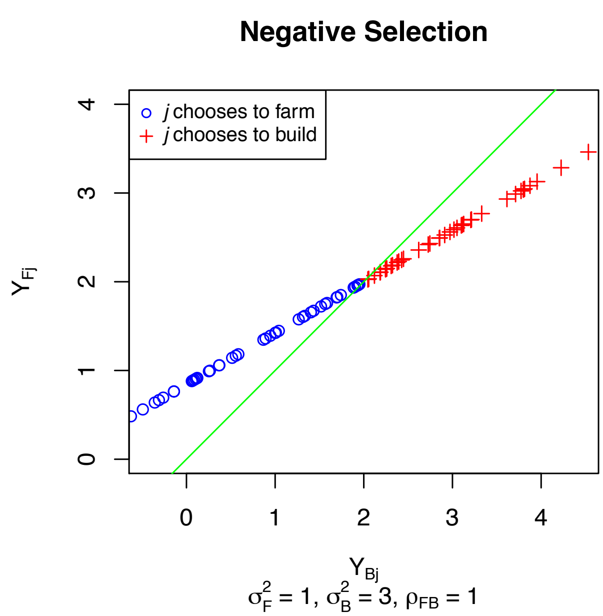
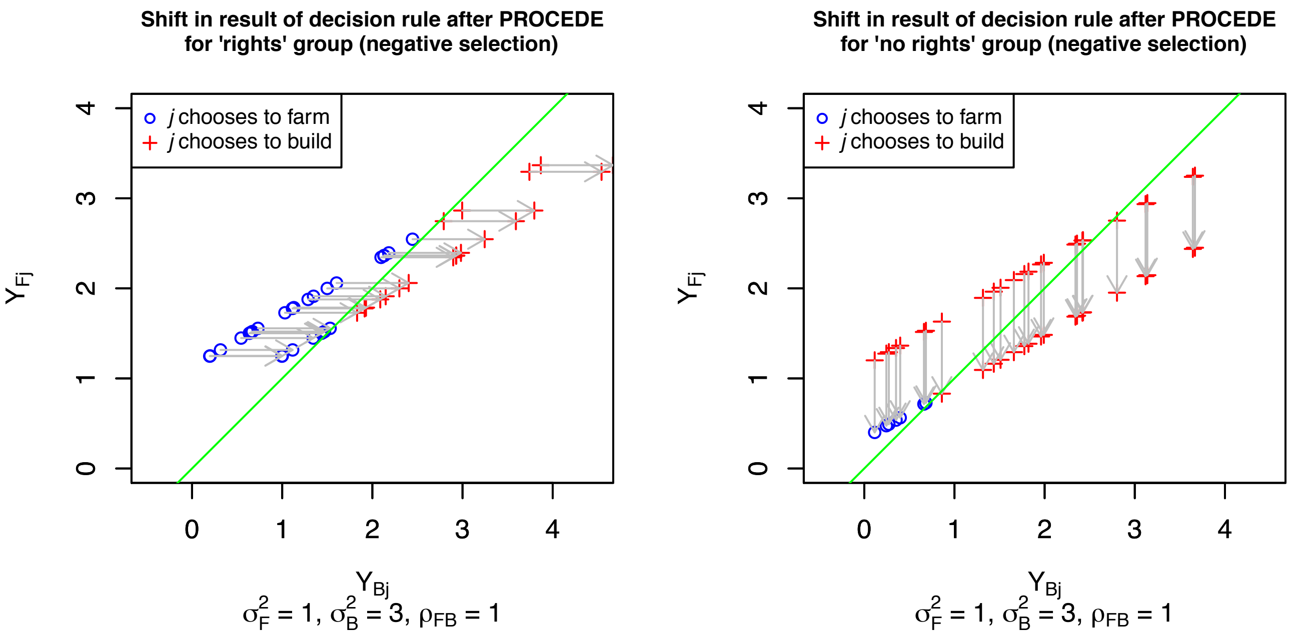


Figure 5



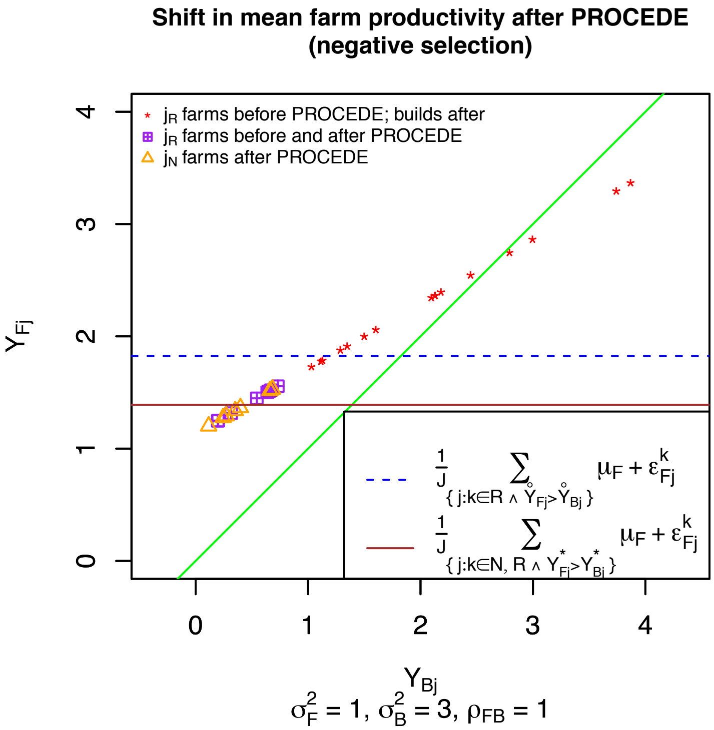
$$\begin{aligned} \overset{*}{Y}{}^N_{Bj} &= \mu_B + \epsilon^N_{Bj} \\ \overset{*}{Y}{}^N_{Fj} &= \mu_F - L + \epsilon^N_{Fj} \end{aligned} \quad , \quad \overset{*}{C}{}^N_j = \begin{cases} B & \text{if } \overset{*}{Y}{}^N_{Bj} > \overset{*}{Y}{}^N_{Fj} \\ F & \text{if } \overset{*}{Y}{}^N_{Bj} \leq \overset{*}{Y}{}^N_{Fj} \end{cases}$$

Earlier I assumed that the random vectors $[\ \epsilon^R_{Bj} \quad \epsilon^R_{Fj}\ ]$ and $[\ \epsilon^N_{Bj} \quad \epsilon^N_{Fj}\ ]$ are identically distributed. This assumption is somewhat unrealistic. The no rights group may be systematically less skilled at agriculture as they are unfamiliar with agricultural techniques. In other words, the mean of ϵ^N_{Fj} may be less than zero. Modification of the model to accommodate this possibility does not change the results I discuss below, since letting the mean of ϵ^N_{Fj} be negative is tantamount to increasing the magnitude of L.

The shifts in the value $\mathring{Y}^R_{Bj}$ to $\overset{*}{Y}{}^R_{Bj}$ and $\mathring{Y}^N_{Fj}$ to $\overset{*}{Y}{}^N_{Bj}$ under conditions of negative selection into farming are illustrated in Figure 5. The Y_{Bj} value of the "rights" group rises. Those agents that are close to the margin move across the 45-degree line. It is only those agents at the lower end of the distribution that choose to stay in farming. The accompanying graph shows that the Y_{Fj} value of the "no rights" group falls. For the no rights group, again the agents close to the decision margin move across the decision line and the only agents remaining to the left of the 45-degree line are those with farming skill much below average.

Figure 6 shows the three groups of agents who ever engaged in farming: those with usufruct rights who chose to leave farming after PROCEDE, those with usufruct rights who chose to continue farming after

Figure 6



PROCEDE, and those without usufruct rights who chose to start farming after PROCEDE. The figure also displays horizontal lines that represent the means of Y_{Fj} given that $C_j = F$. The downward shift in this line indicates that technical efficiency has fallen after PROCEDE under conditions of negative selection. Y_{Fj} is not exactly technical efficiency u_j, but Y_{Fj} is an affine transformation (given by (4)) of u_j, which preserves ranking. Therefore, mean technical efficiency will have fallen if $E\left[Y_{Fj}|C_j = F\right]$ has declined.

With graphs I have illustrated that the rightholders' sale option unambiguously intensifies negative selection into farming. In the appendix I prove this claim formally and also deal with the cases of positive selection and refugee selection. When positive selection or refugee selection prevail, implementation of PROCEDE leads to a rise in mean technical efficiency in agriculture. These results still hold if the assumption that ϵ_{Bj} and ϵ_{Fj} are distributed jointly normal is relaxed. The only condition needed is that the difference between ϵ_{Bj} and ϵ_{Fj} has a log-concave distribution. The class of log-concave distributions includes the normal, exponential, uniform, extreme value, and Laplace distributions, *inter alia*.

Despite the fact that mean technical efficiency in agriculture falls when negative selection prevails, the mean level of skills employed in the economy as a whole unambiguously rises when $\mu_F = \mu_B$, as I show in the appendix. GDP rises since total income increases as the level of employed skills increases. This result is consistent with standard economic intuition: economy-wide efficiency should rise when transactions barriers are lowered. The condition $\mu_F = \mu_B$ is chosen to achieve a closed-form expression for the rise in economy-wide efficiency. The magnitude of the efficiency gain is a function of the variances of ϵ_{Bj} and ϵ_{Fj} and their covariance.

5 Data source & forming the aggregate panel

To estimate technical efficiency, I use the 1991 and 2007 Mexican agricultural censuses. No other agricultural census was carried out in the intervening years. The censuses collected data from every agricultural producer.

Individual records from the 1991 census are not linked to the 2007 census, so these datasets do not constitute a true panel. I constructed a pseudo-panel along the lines of Deaton (1985) whereby the unit of analysis will become the ejido rather than the farm. The outputs and inputs of each farm were summed into an ejido aggregate. This may seem to be an unhappy compromise, but forming a pseudo-panel actually aids the identification strategy.

A traditional panel is vulnerable to bias that arises when respondents drop out of the survey in a nonrandom manner. A pseudo-panel can help eliminate attrition bias since all individuals who qualify for sampling have an equal probability of being sampled in each survey round. In my application, attrition is, in a sense, the object of interest. The ejido as a whole becomes more or less efficient because individuals are more free to select into or out of farming due to PROCEDE. In this line of thinking, the production unit is the ejido rather than the individual farms that comprise the ejido.

The censuses question farmers about the amount of land under cultivation, amount of irrigated land, and the number of workers. For all other inputs, the questions were worded as binary yes-or-no. To handle this issue, I considered all land managed by a farmer who answered "yes" to a question on input use as the "input quantity" for that input. Thus the ejido-level aggregate of the inputs denote the quantity of land within the ejido that was managed by someone who used a given input. The inputs are: total land in hectares, irrigated land, improved seeds, chemical fertilizer, organic fertilizer, pesticides and herbicides, animal traction, tractors, technical assistance, and number of people working on the farm.

I examine the production of both maize and a crop value aggregate. Maize is the dominant crop in Mexico. Over 85 percent of farmers in the sample grew maize, so it deserves to be analyzed separately from other crops. Using a combination of output prices reported in the 2007 agricultural census and 1991 prices recorded by the Food & Agriculture Organization, I computed a Fisher value index for every farm in the sample, then applied the Eltetö-Köves-Szulc transitivity correction to ensure the output of all farms was comparable (Diewert 1999). The value index yields a measure of the agricultural output of the ejido sector across all crops.

Respondents' locations are listed by locality in the agricultural census data. I linked localities to ejidos by measuring the distance between the locality points and the polygons that define each ejido. If a locality was located within an ejido's polygon or was within one kilometer of the polygon, the locality was linked to that particular ejido. Some ejidos could not be reliably linked to localities with this method and were therefore dropped from the analysis. Given that they were more remote, the ejidos that were dropped from the sample had only about half as many members as ejidos in the sample on average and tended to be much smaller in land area. Since the land market would be thinner in smaller ejidos, we may expect a productivity change of smaller magnitude in the excluded ejidos, so dropping them from the sample may bias the estimate of the mean productivity impact away from zero. The mean date of PROCEDE implementation, however, is roughly equal for the in- and out-of-sample groups.

6 The data generating process: inputs and outputs

In this section I will outline the process that generates the relationship between outputs and inputs. The main estimating equation will mimic this relationship.

Since Marschak & Andrews (1944), economists have known that direct estimation of the parameters of the primal production function poses an identification challenge. I will review the problem here.

Let there be heterogeneity in technical efficiency across firms. Ignoring time for now, represent this heterogeneity by a scalar random variable Ω_j and define the production relation as

$$y_j = f(\mathbf{x_j}, \boldsymbol{\beta}) \cdot \exp\{\Omega_j\} \tag{5}$$

where

y_j is output of the j'th firm,

$f(\cdot)$ is the production function,

$\mathbf{x_j}$ is a vector of inputs to the production process, and

$\boldsymbol{\beta}$ is a vector of technological parameters common to all firms.

Let all firms exactly know their own value of Ω_j. Since Ω_j is a known component of each firm's production function, it directly influences the firm's optimization behavior, which includes the choice of inputs $\mathbf{x_j}$. Hence, the elements of $\mathbf{x_j}$ are correlated with the Ω_j term. If the econometrician has no or limited knowledge of Ω_j, naive estimation of equation 5 via regression will suffer from endogeneity bias and lead to inconsistent estimation of the technological parameters $\boldsymbol{\beta}$. This productivity transmission bias can be partially addressed via a fixed-effects model, which is explored below.

Decomposition of the disturbance term Ω

Turning to our case of agricultural production of farms in Mexican ejidos, let the production relation be

$$y_{jvt} = f(\mathbf{x_{jvt}}, \boldsymbol{\beta}) + \Omega_{jvt}$$

where

y_{jvt} is yield, in units of output per hectare, of the j'th farm in the v'th ejido in year t

$f(\cdot)$ is the production function,

$\mathbf{x_{jvt}}$ is a vector of inputs to the production process,

$\boldsymbol{\beta}$ is a vector of technological parameters common to all farms, and

Ω_{jvt} is a random variable.

Define $\vee_{vt}$ as the set of farms in the v'th ejido in year t. At this point, assume that farms do not enter or leave ejidos over time, i.e. $\vee_{vt} = \vee_v$, $\forall t$. I will use this notation later.

Decompose Ω_{jvt} in the following way:

$\Omega_{jvt} = \epsilon_{jvt} + \psi_{jv} + \omega_{jvt} + \zeta_t + u_v + \eta_{vt}$

Then the production relation can be re-written as

$$y_{jvt} = f(\mathbf{x_{jvt}}) + \epsilon_{jvt} + \psi_{jv} + \omega_{jvt} + \zeta_t + u_v + \eta_{vt}$$

I define these components of Ω_{jvt} below.

ϵ_{jvt} represents a disturbance term observed by neither the econometrician nor the farmer at the time input levels are chosen. It can be considered any farm- and time-specific shock that occurs too late in the production process for the farm to react or — in the estimation framework — any measurement error on y_{jvt}. In other words, this is a "pure" random noise component that does not give rise to any endogeneity concerns.

ψ_{jv} is a time-invariant random variable particular to the j'th farm. Farmers know this term and incorporate this knowledge into their optimization decisions. Inherent skill or past educational investments would be included in this term.

ω_{jvt} is a random variable particular to the j'th farm and the t'th year. Farmers know this term, or at least can predict it with some accuracy, before the choice of inputs. The term could represent a shock that varies over both time and farm. Learning a new agricultural technique, for example, could be a positive shock to ω_{jvt}. Transient illness could represent a negative shock.

u_v is a random variable particular to each ejido that does not vary over time. Time-invariant soil quality and altitude, among other factors, can motivate this term. Farmers know the value of u_v and take it into account when choosing input levels.

ζ_t is a shock common to all farmers in Mexico at time t. This can represent shifts in agricultural technology over time. Farmers also know this shock.

η_{vt} is a shock inherent to a particular ejido that varies over time. This term may consist of weather and natural disasters. Farmers know this shock at the time of input choice.

Now assume that $f(\cdot)$ takes a quadratic form. Since most of the agricultural inputs in the data are inessential, a Cobb-Douglas or translog form would be inappropriate. I choose the quadratic form because it permits positive output when some inputs are zero, and represents a flexible functional form. Therefore, the production relation is

$$y_{jvt} = \beta_0 + \sum_m \beta_j X_{jvt}^m + \sum_m \sum_n \beta_{ij} X_{jvt}^m X_{jvt}^n + \epsilon_{jvt} + \psi_{jv} + \omega_{jvt} + \zeta_t + u_v + \eta_{vt}$$

where

X_{jvt}^m is the level of the m'th input chosen by the j'th farm in the v'th ejido in the t'th year,

$\boldsymbol{\beta}$ is a vector of technology parameters, and

all other symbols are defined as previously.

Aggregation to the ejido level

Compute the mean of yield across farms in each ejido and form a vector composed of the aggregates:

$$\mathbf{y_t} = \left[\begin{array}{ccc} \frac{1}{J_1} \sum_{j \in \vee_1} y_{j1t} & \cdots & \frac{1}{J_V} \sum_{j \in \vee_V} y_{jVt} \end{array} \right]$$

where $\mathbf{y_t}$ is a $1 \times V$ vector and J_v is the number of farmers in the v'th ejido. Let y_{vt} represent the v'th element.

Such aggregation then can be used to define an ejido-level production function:

$$\frac{1}{J_v} \sum_{j \in \vee_v} y_{jvt} = \frac{1}{J_v} \sum_{j \in \vee_v} \left(\beta_0 + \sum_m \beta_j X^m_{jvt} + \sum_m \sum_n \beta_{ij} X^m_{jvt} X^n_{jvt} + \epsilon_{jvt} + \psi_{jv} + \omega_{jvt} + \zeta_t + u_v + \eta_{vt} \right) \quad (6)$$

Let $\frac{1}{J_v} \sum_{j \in \vee_v} y_{jvt} = \overline{y_{jvt}}$. Denote the averages of the other variables similarly. Re-express (6) as

$$\overline{y_{jvt}} = \beta_0 + \sum_m \beta_j \overline{X^m_{jvt}} + \sum_m \sum_n \beta_{ij} \overline{X^m_{jvt} X^n_{jvt}} + \overline{\epsilon_{jvt}} + \overline{\psi_{jv}} + \overline{\omega_{jvt}} + \zeta_t + u_v + \eta_{vt}$$

Now assume that $\overline{\omega_{jvt}} = 0$. Time- and farm-specific shocks are challenging to handle, especially with panel data collected years apart. Aggregation at the ejido level will attenuate the impact of these shocks. The data generating process under this assumption is therefore

$$\overline{y_{jvt}} = \beta_0 + \sum_m \beta_j \overline{X^m_{jvt}} + \sum_m \sum_n \beta_{ij} \overline{X^m_{jvt} X^n_{jvt}} + \overline{\epsilon_{jvt}} + \overline{\psi_{jv}} + \zeta_t + u_v + \eta_{vt} \quad (7)$$

Now remove the assumption that $\vee_{vt} = \vee_v, \forall t$. This permits "churning" in the farmer composition of each ejido.

Neither u_v nor η_{vt} change because they are particular to ejidos, not farms.

$\overline{\epsilon_{jvt}}$ does not change because it is i.i.d. across farms, ejidos, and time.

Let $\kappa_{1991-2007}$ represent technological change over 1991-2007. Then ζ_t can be decomposed into

$$\zeta_t = \kappa_{1991-2007} + \sum_{v=1}^{V} \left[\frac{1}{J_{v,2007}} \sum_{j \in \vee_{v,2007}} \psi_{jv} - \frac{1}{J_{v,1991}} \sum_{j \in \vee_{v,1991}} \psi_{jv} \right]$$

So ζ_t represents the sum of any technological improvement and the average change across all ejidos in the skill composition of ejidos. The second term will be nonzero if there is some sort of general time trend of selection into or out of farming based on technical efficiency.

And $\overline{\psi_{jv}}$ gains a t subscript, becoming

$$\overline{\psi_{jvt}} = \frac{1}{J_{vt}} \sum_{j \in \vee_{vt}} \psi_{jv}$$

If positive selection or refugee selection in the Roy framework prevails, then we should expect that

$$\frac{1}{|\mathscr{P}_1|} \sum_{v \in \mathscr{P}_1} \overline{\psi_{jvt}} \quad > \quad \frac{1}{|\mathscr{P}_0|} \sum_{v \in \mathscr{P}_0} \overline{\psi_{jvt}}$$

where $\mathscr{P}_1$ is the set of ejidos where PROCEDE was implemented and $\mathscr{P}_0$ is the set of ejidos where it was not implemented. $|\cdot|$ denotes number of elements of the set. If negative selection prevails, then we will have

$$\frac{1}{|\mathscr{P}_1|} \sum_{v \in \mathscr{P}_1} \overline{\psi_{jvt}} \quad < \quad \frac{1}{|\mathscr{P}_0|} \sum_{v \in \mathscr{P}_0} \overline{\psi_{jvt}}$$

7 Production function: main estimating equation

To identify (7) I estimate

$$\overline{y_{jvt}} = \hat{\beta}_0 + \sum_m \hat{\beta}_j \overline{X^m_{jvt}} + \sum_m \sum_n \hat{\beta}_{ij} \overline{X^m_{jvt} X^n_{jvt}} + \hat{\lambda}_t + \hat{\delta}_v + H(\hat{\boldsymbol{\theta}}, \mathbf{w_{vt}}) + \hat{\gamma} \cdot PROCEDE_{vt} + \overline{\epsilon_{jvt}} \tag{8}$$

The time fixed effects λ_t absorb the ζ_t efficiency disturbance term.

Ejido fixed effects δ_v absorb the u_v efficiency disturbance term and the mean of the $\overline{\psi_{jvt}}$ term.

$H(\hat{\boldsymbol{\theta}}, \mathbf{w_{vt}})$ represents precipitation, solar radiation, and growing degree days, and their squares, in the spring-summer growing season. These variables are also interacted with amount of land with irrigation. This approach matches that of Deschênes & Greenstone (2007). The weather data is derived from the AgMERRA climate dataset (Ruane, Goldberg, & Chryssanthacopoulos 2015). If the variation in the η_{vt} disturbance term is due only to these weather variables, including $H(\hat{\boldsymbol{\theta}}, \mathbf{w_{vt}})$ in the regression removes any endogeneity issues associated with η_{vt}.

Finally let

$$PROCEDE_{vt} = \sum_{z=1993}^{t} \mathbb{1}\left\{v \in \mathscr{P}_{1z}\right\}$$

That is to say, $PROCEDE_{vt}$ is the cumulative number of years that a given ejido has been PROCEDE-certified. Here $\hat{\gamma}$ measures the marginal (linear) effect of one additional year of ejido members having land titles. Among ejidos that implemented PROCEDE, by 2007 an average of 9 years had elapsed since implementation. This average of 9 years is important for interpretation of the coefficient magnitudes in the following section. The coefficients can be scaled by 9 to approximate the effect of PROCEDE on the average ejido by 2007.

In the analysis I omit ejidos that never implemented PROCEDE. Defining $PROCEDE_{vt}$ in this way avoids relying on identification of the effect of PROCEDE based on the only 9 percent of ejidos that did not implement PROCEDE. There is also a practical data issue with including ejidos that did not implement PROCEDE: these ejidos do not have a GIS polygon since they were never mapped as part of the PROCEDE program. Their locations are known with no greater specificity than their municipality, which is roughly equivalent to a U.S. county. Therefore, linking them to localities would be subject to much greater error.

8 Results

Table 1 displays the OLS estimate of (8). The estimated effect of PROCEDE on the technical efficiency of maize production and crops as a whole is not statistically different from zero. This result is consistent with the Roy theoretical framework, since change of land ownership resulting from differences in comparative advantage may not shift the average level of absolute advantage among farmers in post-PROCEDE ejidos. A small proportion of ejidos known as "agrarian communities" were originally formed out of pre-existing communities and therefore could have had a different response to PROCEDE. However, the null hypothesis of no effect still cannot be rejected when these communities are excluded from the sample.

Identification of the effect of PROCEDE here relies on some time lag between implementation of PROCEDE and achievement of the new efficiency equilibrium. The signal of the effect of PROCEDE may be obscured if the transition to the new equilibrium was rapid. To investigate this possibility I estimated a number of specifications that transformed the treatment variable nonlinearly (not shown), but the null hypothesis of no effect could still not be rejected. The discussion below on the evidence for effects on land transactions help address to the question of the time lag between PROCEDE implementation and transition to the new equilibrium.

Table 1: Effect of PROCEDE on efficiency

	Dependent variable:	
	Mean maize yield in metric tons per ha	Mean value index yield in index points per ha
Years since PROCEDE	0.018	0.039
	(0.012)	(0.035)
Dependent variable mean	1.63	2.65
Dependent variable st. dev.	1.74	4.67
Observations	33,154	35,340
Adjusted R^2	0.639	0.351

Note: *p<0.1; **p<0.05; ***p<0.01
Technological and weather parameters omitted from table
Ejido and state-by-year fixed effects
Standard errors, in parentheses, clustered at state level

Full property rights

Embedded within PROCEDE was an option to go one step further in land tenure: full property rights (*dominio pleno*). The 1992 reforms allowed, in all ejidos, unrestricted rent of ejido land but prohibited sale to anyone who was not a member of the same ejido. Members of an ejido could democratically decide to allow the sale of land to individuals who were not members of the ejido. About 16 percent of ejidos opted for full property rights.

Analyzing the impact of full property rights has some benefits and drawbacks compared with measuring the effect of PROCEDE. Identification of the effect of PROCEDE relied upon an assumption that the efficiency effects were not immediate. If the full effects of PROCEDE were realized the year they were implemented, any effects would have been undetectable. The full property rights regime, on the other hand, was implemented in some ejidos and left completely unimplemented in others.

Implementation of full property rights may threaten the identification assumptions of the fixed-effects model, however. The timing of PROCEDE implementation was partly controlled by the central government. In contrast, ejido members had greater control over whether full property rights would be implemented in their ejido. Compared with agents from the central government who were not as well informed about local conditions, they would have been more likely to choose the full property rights regime in response to some ejido-specific shock that was associated with agricultural efficiency. Endogeneity bias may be present. A birthrate boom is one possible shock where endogeneity concerns may arise. Younger and stronger farmers may be more efficient, while at the same time the community may be more inclined to implement full property rights to make it easier to securely transfer title to the larger number of heirs. In this scenario the estimated effect of full rights would be biased upward.

Table 2 displays the estimated effect of the full rights regime on agricultural efficiency, the counterpart to Table 1. The estimated improvement in efficiency in maize is economically meaningful at a rise of about one third of a metric ton per hectare, and statistically significant at the one percent level. The estimated effect in the crop sector as a whole is about 14 percent of the standard deviation of the dependent variable, and it is statistically significant at the one percent level.

Table 2: Effect of full property rights on efficiency

	Dependent variable:	
	Mean maize yield in metric tons per ha	Mean value index yield in index points per ha
Full rights	0.314***	0.663***
	(0.100)	(0.196)
Dependent variable mean	1.63	2.65
Dependent variable st. dev.	1.74	4.67
Observations	33,154	35,340
Adjusted R^2	0.639	0.351

Note: *p<0.1; **p<0.05; ***p<0.01
Technological and weather parameters omitted from table
Ejido and state-by-year fixed effects
Standard errors, in parentheses, clustered at state level

The results here do not come without a further caveat. According to anecdotal evidence, one factor motivating ejidos to adopt the full property regime was that it enabled the sale of property for the purposes of peri-urban real estate development. If a large amount of low-productivity land was taken out of agriculture to feed urban growth, we would observe a rise in average agricultural productivity in the ejido. This rise would have nothing to do with the hypothesized shift in average farmer skill under study in this paper. On the other hand, proximity to urban areas would facilitate certain individuals pursuing their comparative advantage and shifting to off-farm employment, accounting for greater efficiency effects. The tools and data available do not allow separate examination of the issues of urban development and shifts in occupation type.

The Mexican government maintains digitized maps of the perimeters of settlements with greater than 15,000 inhabitants. This dataset allows the empirical exploration of the urban proximity issue. In fact there is a strong relationship between an ejido's distance to nearest urban area and implementation of full property rights. The correlation between adoption of full rights and the log of distance to the nearest urban area is -0.39, seemingly confirming the anecdotal evidence.

Table 3: Heterogeneous effect of full property rights according to distance from urban area

	Dependent variable:			
	Maize	Index	Maize	Index
Full rights	0.517***	1.014***	0.183	0.407*
	(0.183)	(0.340)	(0.145)	(0.225)
Full rights × $sinh^{-1}$(Dist to urban area)	−0.117*	−0.198		
	(0.071)	(0.133)		
Full rights × Dist to urban area ≤ median			0.256	0.515
			(0.245)	(0.343)
Observations	33,154	35,340	33,154	35,340
R^2	0.821	0.678	0.821	0.678
Adjusted R^2	0.640	0.352	0.639	0.352

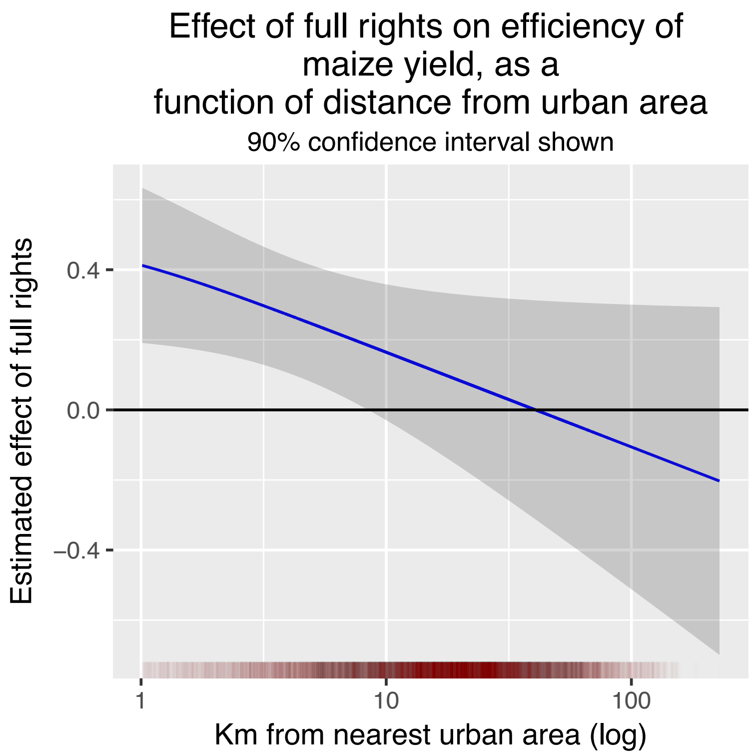
Note: *p<0.1; **p<0.05; ***p<0.01
Median distance is median conditional on having full rights
Technological and weather parameters omitted from table
Ejido and state-by-year fixed effects
Standard errors, in parentheses, clustered at state level

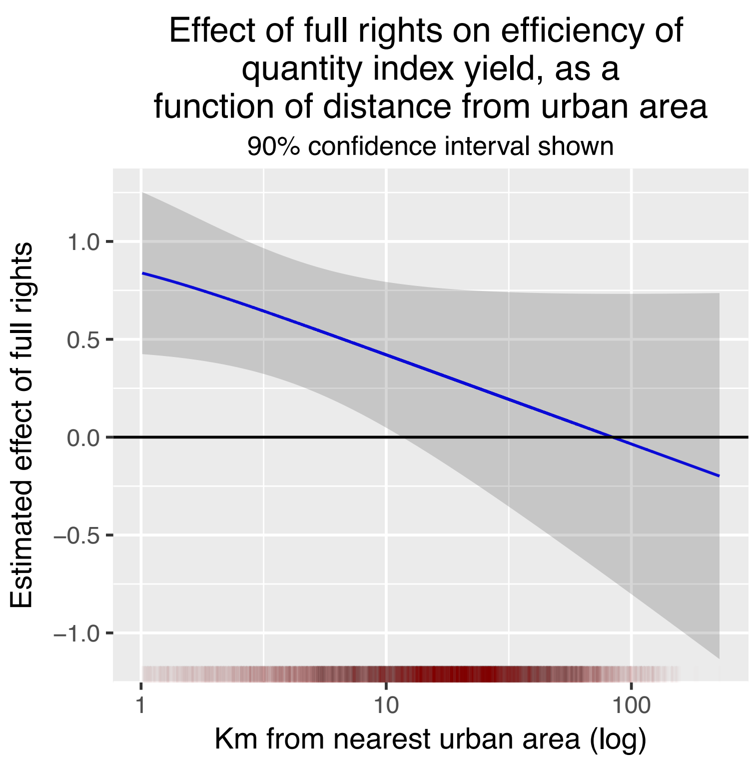
Interacting the full rights regime dummy with distance to urban center can reveal whether the positive effect on efficiency only occurs when ejidos are near urban areas. Table 3 indicates that the estimated effect of full rights on maize efficiency does not achieve conventional levels of significance for ejidos that are further away from urban centers than the median ejido. Estimates for the interaction with distance are also included to illustrate the distance threshold at which the effect becomes not statistically significant. Distance is transformed by inverse hyperbolic sine, which is similar to a log transformation but is defined at zero. The estimated effect as a function of distance from urban area is plotted in Figure 9. The 90 percent confidence interval for the effect of full rights on efficiency crosses zero when the distance from nearest urban area is greater than 10 kilometers.

Heterogenous effects

The theoretical model predicts ambiguous effects of land titling. The sign of the effect depends on the values of the elements of the covariance matrix for the distributions of skill-driven outcomes in the agricultural and off-farm sectors. Say that the variance of skill in off-farm activity takes a common value across the

Figure 7


Effect of full rights on efficiency of maize yield, as a function of distance from urban area
90% confidence interval shown
Estimated effect of full rights
0.4
0.0
−0.4
1
10
100
Km from nearest urban area (log)


Effect of full rights on efficiency of quantity index yield, as a function of distance from urban area
90% confidence interval shown
Estimated effect of full rights
1.0
0.5
0.0
−0.5
−1.0
1
10
100
Km from nearest urban area (log)

country. Such an assumption is roughly true if regional labor markets in Mexico are integrated. Also set the correlation so that only positive selection or negative selection, not refugee selection, can occur. Then regions with higher variance in agricultural skill will see positive sorting into agriculture, and land titling would result in a rise in efficiency. An obvious source of different variance in agricultural outcomes is distinct agroclimactic conditions.

Table 4: Heterogeneous effects of PROCEDE by bioregion

	Dependent variable:	
	Maize yield	Value index yield
Yrs since PROCEDE × Altiplano Norte	-0.064^{***} (0.014)	-0.021 (0.099)
Yrs since PROCEDE × Altiplano Sur	0.005 (0.014)	-0.032 (0.055)
Yrs since PROCEDE × Baja California	-0.093 (0.356)	0.276 (0.613)
Yrs since PROCEDE × California	-0.220 (0.101)	-1.303^{*} (0.459)
Yrs since PROCEDE × Costa del Pacifico	-0.002 (0.026)	0.053 (0.049)
Yrs since PROCEDE × Del Cabo	-0.438^{***} (0.050)	1.610^{***} (0.194)
Yrs since PROCEDE × Depresion del Balsas	0.048 (0.023)	0.344 (0.395)
Yrs since PROCEDE × Eje Volcanico	0.030 (0.047)	0.042 (0.037)
Yrs since PROCEDE × Golfo de Mexico	-0.014 (0.017)	0.007 (0.043)
Yrs since PROCEDE × Los Altos de Chiapas	0.032^{***} (0.005)	0.008 (0.015)
Yrs since PROCEDE × Oaxaca	0.013 (0.023)	-0.004 (0.027)
Yrs since PROCEDE × Peten	0.017 (0.029)	0.003 (0.013)
Yrs since PROCEDE × Sierra Madre del Sur	-0.007 (0.024)	0.062 (0.066)
Yrs since PROCEDE × Sierra Madre Occidental	0.031 (0.016)	0.140 (0.052)
Yrs since PROCEDE × Sierra Madre Oriental	0.011 (0.010)	0.006 (0.039)
Yrs since PROCEDE × Soconusco	-0.011 (0.012)	-0.026 (0.013)
Yrs since PROCEDE × Sonorense	0.056 (0.082)	-0.091 (0.248)
Yrs since PROCEDE × Tamaulipeca	-0.028 (0.013)	0.021 (0.058)
Yrs since PROCEDE × Yucatan	0.016 (0.034)	0.068 (0.072)
Wald stat of all coefficients are equal	358.5^{***} (df = 18)	815.8^{***} (df = 18)
Wald stat of all coefficients are zero	626.5^{***} (df = 19)	912^{***} (df = 19)
Observations	33,112	35,282
R^2	0.813	0.675
Adjusted R^2	0.623	0.346
Note:		*p<0.1; **p<0.05; ***p<0.01
		Tech and weather parameters omitted from table
		Ejido and bioregion-by-year fixed effects
		Standard errors clustered at state level
		p-value flags adjusted for multiple hypotheses

Tables 4 and 5 display the results of estimating the heterogenous effects of PROCEDE and full property rights. Full property rights and years elapsed since PROCEDE are interacted with dummies for Mexican biogeographic regions. Given that multiple hypotheses are being tested, the p-value star thresholds were adjusted to ensure that the star flags reflect the family-wise type I error rate in each regression using the

Table 5: Heterogeneous effects of full rights by bioregion

	Dependent variable:	
	Maize yield	Value index yield
Full rights × Altiplano Norte	0.615 (0.432)	0.797 (0.947)
Full rights × Altiplano Sur	0.408 (0.264)	0.252 (0.559)
Full rights × Baja California		3.248 (9.757)
Full rights × California	−2.673 (1.176)	3.201*** (0.633)
Full rights × Costa del Pacifico	0.183 (0.328)	0.842 (0.776)
Full rights × Del Cabo	−1.741*** (0.492)	−2.880*** (0.817)
Full rights × Depresion del Balsas	0.868 (0.387)	8.753 (8.515)
Full rights × Eje Volcanico	0.522* (0.182)	0.803 (0.509)
Full rights × Golfo de Mexico	−0.085 (0.132)	0.285 (0.238)
Full rights × Los Altos de Chiapas	−0.383*** (0.091)	−0.418*** (0.119)
Full rights × Oaxaca	0.418 (0.225)	1.200** (0.366)
Full rights × Peten	−0.277*** (0.071)	−0.383 (0.230)
Full rights × Sierra Madre del Sur	0.260 (0.415)	2.324*** (0.408)
Full rights × Sierra Madre Occidental	−0.015 (0.210)	2.007 (0.915)
Full rights × Sierra Madre Oriental	0.424 (0.174)	0.536* (0.191)
Full rights × Sonorense	0.529 (0.500)	0.579 (1.088)
Full rights × Tamaulipeca	0.195 (0.364)	0.880 (0.892)
Full rights × Yucatan	−0.119 (0.119)	−0.495 (0.338)
Wald stat of all coefficients are equal	441.7*** (df = 16)	369.6*** (df = 17)
Wald stat of all coefficients are zero	453.5*** (df = 17)	442.5*** (df = 18)
Observations	33,112	35,282
R^2	0.814	0.678
Adjusted R^2	0.626	0.353

Note: *p<0.1; **p<0.05; ***p<0.01
Technological and weather parameters omitted from table
Ejido and bioregion-by-year fixed effects
Standard errors, in parentheses, clustered at state level
p-value flags adjusted for multiple hypothesis testing

method described by Hothorn, Bretz, & Westfall (2008). PROCEDE is estimated to have boosted maize efficiency in two regions (Altiplano Norte and Los Altos de Chiapas), but reduced efficiency in one of them (Del Cabo). Furthermore, Wald tests strongly reject the null hypotheses that all coefficients are equal to each other and that all coefficients are zero. Hence, PROCEDE had heterogenous effects on efficiency in different biogeographic regions. This result, combined with the theoretical model, suggests that the variance of agricultural outcomes was higher in some biogeographic regions than others.

Effect on land transactions and land concentration

Indirect evidence from PROCAMPO recipient data

PROCAMPO is a subsidy program initiated in 1994 that was intended to soften the blow of sharply lower grain prices following the implementation of NAFTA. PROCAMPO granted a transfer to farmers based on the amount of land under cultivation. Analysis of the public data on PROCAMPO disbursements, which includes name of beneficiary, name of ejido, and amount of land in the program, can shed light on land transactions patterns.

To be eligible for enrollment in PROCAMPO, a given plot of land had to be planted with one of nine staples in the 1991-1993 period (Sadoulet et al. 2001). Land parcels were grandfathered into the program based on this criteria. To maintain eligibility for a parcel each season, landowners are required to use the land for some productive purpose such as agriculture, livestock grazing, or forest activity (Cord & Wodon 2001). Subsidy amounts have ranged from about 70 to 100 USD per hectare over the course of the program (Cord & Wodon 2001, USDA 2013). The enrollment barrier being low and benefits being relatively high, about 84 percent of ejido members participate, while 90 percent of Mexico's total cultivated area is in the program (Cord & Wodon 2001). Thus the PROCAMPO recipient data approaches universal coverage for land cultivation dynamics at the farmer level.

The Mexican agricultural ministry maintains a database of over 50 million records of PROCAMPO payments. Via probabilistic string matching, I was able to link these records to about 85 percent of the ejidos in PHINA, the definitive inventory of ejidos in Mexico. My universe is the 1995-2012 spring-summer growing seasons.

PROCAMPO payments follow the user of the land, so a renter or purchaser of parcels appears as a beneficiary in the PROCAMPO database (Cord & Wodon 2001). Therefore, individual land tenure patterns can be detected via linking individuals over time. The dataset does not allow transactions to be directly observed, but it does quantify the rise and fall of landholdings for each individual, from which I might infer aggregate transaction statistics. Individuals may increase or reduce the amount of land in PROCAMPO for a variety of reasons, including choosing not to plant in a given season. My claim is that under certain conditions the effect of PROCEDE on land transactions can be revealed by joint analysis of the rise and fall of landholdings for each individual.

Let S_{jvt} and B_{jvt} be the amount of land sold or rented out and bought or rented in, respectively, by the j'th individual in the v'th ejido in the t'th year. Then the following accounting identity must hold:

$$\sum_{j=1}^{J_v} S_{jvt} = \sum_{j=1}^{J_v} B_{jvt}$$

That is to say, the sum of all land in a given ejido sold or rented out must equal the sum of all land in that ejido that is bought or rented in. Let O^-_{jvt} be the amount of land that an individual removes from PROCAMPO due to reasons other than a transaction. Define O^+_{jvt} similarly. Then define

$$D^-_{jvt} \equiv S_{jvt} + O^-_{jvt}$$

$$D^+_{jvt} \equiv B_{jvt} + O^+_{jvt}$$

Define transactions as $T_{vt} \equiv \sum_{j=1}^{J_v} S_{jvt} \equiv \sum_{j=1}^{J_v} B_{jvt}$. Summing within each ejido yields

$$D^-_{vt} = T_{vt} + O^-_{vt}$$

$$D^+_{vt} = T_{vt} + O^+_{vt}$$

D^-_{vt} and D^+_{vt} can be calculated from the PROCAMPO beneficiary data. Denote PROCAMPO landholdings for individual j in ejido v at time t as $L_{j,v,t}$. Then the empirical counterpart to D^+_{vt} is the sum of all individuals' positive shifts in landholding from year to year in a given ejido:

$$D^+_{vt} = \sum_{j\in v} (L_{j,v,t} - L_{j,v,t-1}) \cdot \mathbb{1}\{L_{j,v,t} > L_{j,v,t-1}\}$$

The magnitude of the sum of negative shifts can be defined similarly:

$$D^-_{vt} = \left| \sum_{j\in v} (L_{j,v,t} - L_{j,v,t-1}) \cdot \mathbb{1}\{L_{j,v,t} < L_{j,v,t-1}\} \right|$$

Now introduce P_{vt}, which is a binary variable that takes a value 1 in the year that PROCEDE is introduced in ejido v and every subsequent year. The following relationship can be estimated by OLS:

$$D^-_{vt} = \beta_0 + \beta_{P^-} P_{vt} + \epsilon_{vt}$$

Since this is a regression on an intercept and a single variable, $\hat{\beta_{P^-}}$ is a simple function the variance of P_{vt} and the covariance of D^-_{vt} and P_{vt}:

$$\hat{\beta_{P^-}} = \frac{\hat{\sigma_{PD^-}}}{\hat{\sigma^2_P}} = \frac{\hat{\sigma_{PT}} + \hat{\sigma_{PO^-}}}{\hat{\sigma^2_P}}$$

Regression of D^-_{vt} on an intercept and P_{vt} yields a similar expression:

$$\hat{\beta_{P+}} = \frac{\hat{\sigma_{PT}} + \hat{\sigma_{PO^+}}}{\hat{\sigma_P^2}}$$

Assume that either $\sigma_{PO^-} \leq 0$ or $\sigma_{PO^+} \leq 0$

Then a necessary condition for both β_{P^-} and β_{P^+} to be positive is that $\sigma_{PT} > 0$.

If on the other hand both $\sigma_{PO^-} > 0$ and $\sigma_{PO^+} > 0$, then $\beta_{P^-} > 0$ and $\beta_{P^+} > 0$ implies only $\sigma_{PT} > -\sigma_{PO^-}$ and $\sigma_{PT} > -\sigma_{PO^+}$, leaving the sign of σ_{PT} unknown.

These facts are intuitive. With the data available we cannot distinguish between a situation in which PROCEDE caused a rise in transactions and a situation in which PROCEDE led to both a rise in the quantity of land removed from PROCAMPO and a rise in land that was introduced into PROCAMPO. The latter scenario lacks theoretical motivation. Any factor that increases the attractiveness of putting land into production would lead to a corresponding reduction in the attractiveness of taking land out of production, and vice versa. Therefore if $\sigma_{PO^+} > 0$, we would expect $\sigma_{PO^-} < 0$ or at least $\sigma_{PO^-} = 0$. In such a case rejecting the hypotheses of both $\beta_{P^-} \leq 0$ and $\beta_{P^+} \leq 0$ would be equivalent to rejecting the hypothesis that $\sigma_{PT} \leq 0$. Hence, in the absence of the perverse scenario of $\sigma_{PO^-} > 0$ and $\sigma_{PO^+} > 0$, the sign of σ_{PT} is correctly identified as positive if $\beta_{P^-} > 0$ and $\beta_{P^+} > 0$.

Table 6 displays estimates of β_{P^-} and β_{P^+}. Both parameter estimates are positive and the null hypothesis that they are zero is strongly rejected. D^-_{vt} and D^+_{vt} have been transformed by inverse hyperbolic sine since they are left-censored at zero. An OLS estimate of a censored equation is inconsistent; the estimator is biased toward zero compared with a maximum likelihood tobit estimator (Greene 1981). However, since the tobit model with fixed effects encounters the incidental parameters problem and the object here is to estimate the sign rather than the magnitude of β_{P^-} and β_{P^+}, OLS is satisfactory in this case. As robustness checks I estimate a model in which the left hand side is a discrete count of the number of beneficiaries that changed the amount of land in PROCAMPO and a model with the number of individuals that appear or disappear from the PROCAMPO beneficiary rolls. Results for estimation of these models, displayed in Tables 7 and 8, are also consistent with the hypothesis that PROCEDE increased the rate of land transactions.

Figure 11 displays the estimated values of $\hat{\beta_{P^-}}$ and $\hat{\beta_{P^+}}$ as dummy variables for leads and lags of PROCEDE implementation. Jointly, the values of these coefficients supports the idea there was a burst of transaction activity in the three years following PROCEDE implementation. The temporal pattern has the appearance of a release of pent-up demand for land transactions after PROCEDE was implemented. About four years after implementation of PROCEDE a new landholding equilibrium established itself.

Table 6: Effect of PROCEDE on land transactions

	Dependent variable:	
	ha added to PROCAMPO ($D+$)	ha removed from PROCAMPO ($D-$)
PROCEDE in effect	0.116***	0.162***
	(0.024)	(0.022)
Dependent var mean	2.2	2.25
Dependent var st. dev.	2.09	2.09
Observations	424,386	424,386
R^2	0.732	0.721
Adjusted R^2	0.716	0.704

Note: *p<0.1; **p<0.05; ***p<0.01
Dependent variables transformed by $sinh^{-1}$
Ejido and state-by-year fixed effects
Standard errors, in parentheses, clustered at state level

Table 7: Effect of PROCEDE on land transactions, discrete measure

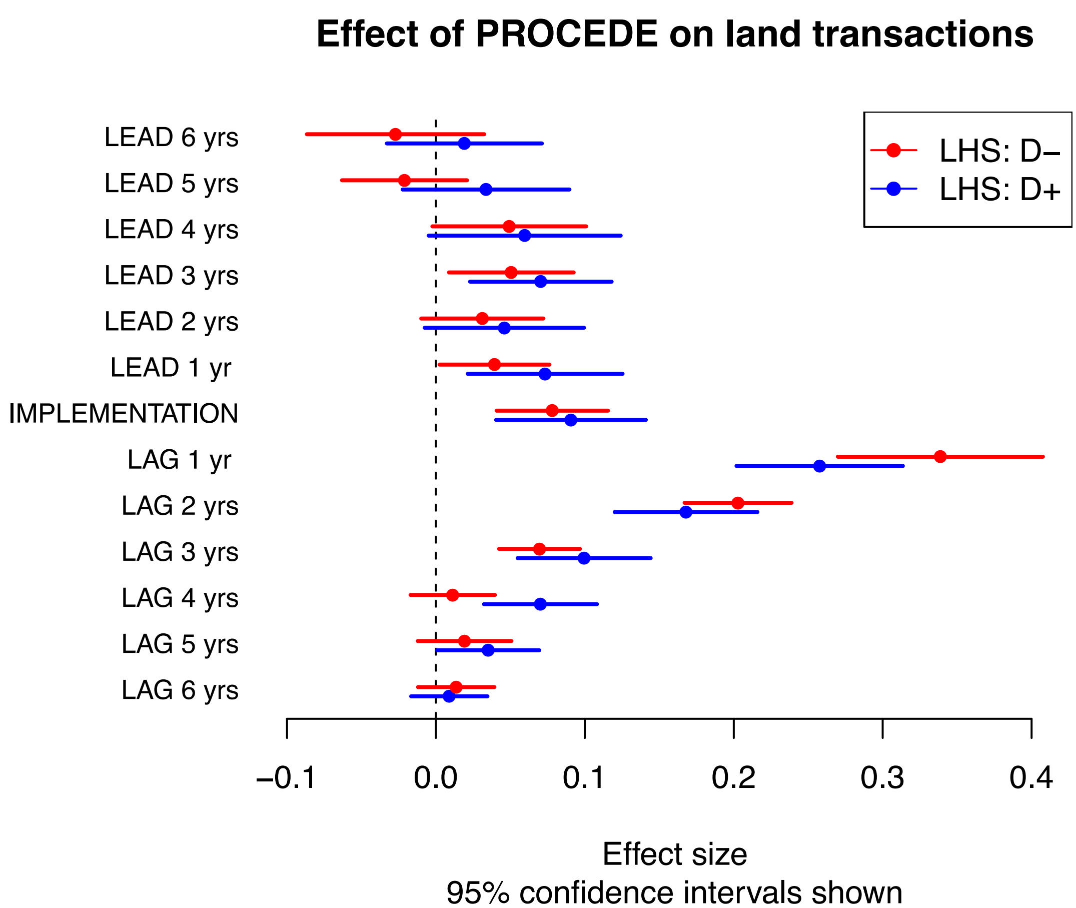
	Dependent variable:	
	# of recipients who added land	# of recipients who removed land
PROCEDE in effect	0.145***	0.193***
	(0.024)	(0.019)
Dependent var mean	1.6	2.25
Dependent var st. dev.	1.58	1.59
Observations	424,386	424,386
R^2	0.753	0.741
Adjusted R^2	0.738	0.725

Note: *p<0.1; **p<0.05; ***p<0.01
Dependent variables transformed by $sinh^{-1}$
Ejido and state-by-year fixed effects
Standard errors, in parentheses, clustered at state level

Table 8: Effect of PROCEDE on land transactions, appear/disappear measure

	Dependent variable:	
	# of recipients newly appearing	# of recipients disappearing
PROCEDE in effect	0.095***	0.127***
	(0.019)	(0.016)
Dependent var mean	1.31	1.41
Dependent var st. dev.	1.48	1.5
Observations	424,386	424,386
R^2	0.733	0.730
Adjusted R^2	0.717	0.714
Note:		*p<0.1; **p<0.05; ***p<0.01
		Dependent variables transformed by $sinh^{-1}$
		Ejido and state-by-year fixed effects
		Standard errors, in parentheses, clustered at state level

Figure 8: Effect of PROCEDE on land transactions



Rent and sharecropping patterns

The 1991 and 2007 agricultural censuses ask respondents for the amount of land under their management by tenure status. Inferring sales is out of reach, but PROCEDE's effects on rental and sharecrop arrangements can be estimated. For this estimation, displayed in Table 9, the unit of observation is the farm rather than the ejido since here there are no concerns about productivity transmission bias.

Table 9: Effect on land tenure

	Dependent variable:			
	Any renting in	Any sharecropping	Hectares rented in	Hectares sharecropped
	Linear probability model		*OLS*	
Years since PROCEDE	0.189***	0.0524***	0.0195***	0.00065
	(0.0178)	(0.0142)	(0.00294)	(0.00208)
Observations	3,768,786	3,768,786	3,768,786	3,768,786
R^2	0.104	0.093	0.039	0.063
Adjusted R^2	0.098	0.086	0.032	0.057

Note: *p<0.1; **p<0.05; ***p<0.01
Linear probability model coefficients expressed as percentage points
Ejido, year, and state-by-year fixed effects
Standard errors, in parentheses, clustered at state level

Farmers living in ejidos that implemented PROCEDE in earlier years were more likely to be renting in land (renting out land was not recorded in the censuses). For every year elapsed since the implementation of PROCEDE, the probability of renting in land rose by about 0.2 percentage points. The estimated effect on the quantity of land rented in was about 0.02 hectares per year elapsed. The estimated effect on sharecropping was positive for the extensive margin, but not statistically different from zero for the intensive margin. These results suggest that PROCEDE encouraged land transactions.

Land inequality and land concentration

The impact of land transactions on land concentration is theoretically ambiguous. The sign of the effect can hinge on the credit market environment, the risk profile for agriculture, and the parameters of the available agricultural technology (Ayala & Liverpool-Tasie 2016). The typical argument points out that decreasing returns to scale makes smaller farms more profitable and therefore more likely to acquire land. In the ejido context, however, the more important consideration may be the floor of the minimum efficient size rather than the ceiling of the inverse productivity-size relationship. Land distribution in ejidos is unusually equal owing to the motivation for creating ejidos. The average land Gini index in ejidos in the pooled 1991 & 2007 sample is 0.37. If there exists a land concentration equilibrium with perfectly working markets, the pre-PROCEDE status quo in ejidos may have been below that equilibrium rather than above it. In such a case, we may expect that land inequality would rise to converge with the equilibrium rather than fall.

Table 10: Effect of PROCEDE on concentration of land under cultivation

	Dependent variable:			
	Land Gini Index		Land Herfindahl Index	
	(1)	(2)	(3)	(4)
Years since PROCEDE	0.00129*		0.00092*	
	(0.00071)		(0.00047)	
2007 year dummy		0.07897***		0.02320***
		(0.00701)		(0.00596)
State-by-year FE	Yes	No	Yes	No
Dependent variable mean	0.3641	0.3641	0.05964	0.05964
Dependent variable st dev	0.1184	0.1184	0.07895	0.07895
Observations	35,436	35,436	35,436	35,436
R^2	0.66844	0.65399	0.67723	0.66175
Adjusted R^2	0.33446	0.30797	0.35209	0.32349

Note: *p<0.1; **p<0.05; ***p<0.01
Ejido fixed effects
Standard errors, in parentheses, clustered at state level

The estimates displayed in Table 10 indicate that PROCEDE led to economically significant rises in both the Gini and Herfindahl indices of land under cultivation, although the point estimates are not very precise. The coefficients for the 2007 year dummy is displayed for comparison purposes. Although related, the Gini and Herfindahl indices deal with distinct distribution concepts. Roughly speaking, the Gini index does not account for differences in the number of landowners. An ejido with 100 hectares divided equally among 50 landowners would have the same Gini index – zero – as one with 100 hectares divided equally among 2 landowners. The Herfindahl index would meanwhile assign a value of 0.02 to this first ejido and 0.5 to the second. Both measures point in the same direction in these regressions, but in the case of the Herfindahl index, PROCEDE accounts for a larger share of the total movement of the index over 1991 - 2007. The magnitude of the coefficient on PROCEDE years elapsed is about 4 percent of the magnitude of the coefficient on the 2007 year dummy in the Herfindahl index, while the corresponding figure for the Gini index is 1.5 percent. Since the average ejido had implemented PROCEDE 9 years prior to 2007, PROCEDE accounted for about 36 percent of the movement in the Herfindahl index in the average ejido over the 1991-2007 period, and for 15 percent of the movement in the Gini index.

Table 11 displays estimates of the impact of PROCEDE on land under cultivation per farm. The unit of observation is the farm. Examining the land concentration question from this perspective yields results similar to the Gini and Herfindahl indices. Every year elapsed since PROCEDE is associated with a modest rise of 200 square meters in farms' cultivated area.

Table 11: Effect of PROCEDE on farm size

	Dependent variable
	Hectares under cultivation per farm
Years since PROCEDE	0.0183**
	(0.00926)
Dependent variable mean	3.989
Observations	3,768,786
Adj. R^2	0.0643
Note:	*p<0.1; **p<0.05; ***p<0.01
	Ejido and state-by-year fixed effects
	Standard error, in parenthesis, clustered at state level

Opponents of PROCEDE and the 1992 ejido reforms feared that liberalization of the land market would lead to re-concentration of land ownership (de Ita 2006). They saw the reforms as reversing the achievements of 1910's Mexican Revolution. The empirical evidence is somewhat consistent with their predictions. Combined with the inability to detect an effect on efficiency, this result is salient for the debate over the tradeoff between efficiency and equality. We do not observe economic outcomes in the off-farm sector, however, so we cannot make a general claim about the economy-wide inequality implications of the empirical results here.

Effect on receipt of credit

The theoretical literature on land titling and the credit market focuses on the potential for collateralizing land once the right to alienate it has been established (Feder & Feeny 1991). In regards to the goal of widening credit access, the 1991 ejido tenure reform snags on its own rules, which permit rent to anyone but restrict sale to only fellow members of a given ejido if the ejido has not opted for full property rights. A commercial bank would be unable to take possession of collateralized land in the event of default. Hence the impact of PROCEDE on the supply side of commercial lending may be limited.

I propose that land titling should also entail shifts in the demand for credit. Positive or negative selection into agriculture (rather than neutral selection) will alter the type of people seeking credit. Whether farmers in the post-PROCEDE equilibrium find themselves demanding greater or lower levels of credit is an open

Table 12: Effect on receipt of credit, extensive margin

	Linear probability model *Dependent variable:*			
	Any credit	Public source	Commercial private source	Other private source
Years since PROCEDE	-0.245***	-0.172***	0.021***	0.0983***
	(0.0625)	(0.0217)	(0.00562)	(0.0140)
Observations	3,768,786	3,768,786	3,768,786	3,768,786
R^2	0.288	0.244	0.054	0.135
Adjusted R^2	0.283	0.239	0.048	0.129

Note: *p<0.1; **p<0.05; ***p<0.01
Coefficients expressed as percentage points
Ejido, year, and state-by-year fixed effects
Standard errors, in parentheses, clustered at state level

question. In an environment of positive selection, high-skill individuals may seek higher levels of agricultural inputs. Higher credit levels may be needed for these inputs, but these individuals may also have access to a deeper pool of liquidity, diminishing the need to seek credit in the first place. One final complication is contextual: withdrawal of the state from involvement in the agricultural sector in Mexico led to dramatic tightening in the credit market between the two census years. The proportion of ejido farmers who received any credit fell from 18 percent to 4 percent in the period under study. Estimation with fixed effects in theory should handle the nationwide shocks, but cautious interpretation is appropriate nevertheless.

Table 12 displays estimates of the effect of PROCEDE on receipt of credit from various sources. Early implementation of PROCEDE is associated with a large fall in the probability of receipt of credit from any source. Receipt of credit from government sources fell by a large magnitude in response to PROCEDE. In contrast, receipt of credit from non-commercial private sources rose modestly; receipt of credit from commercial private sources rose by a small amount.

9 Conclusion

PROCEDE was a large-scale land titling program in Mexico. At the time of implementation, the government hoped to boost the efficiency of agricultural production by liberalizing the land market and placing land into the hands of high-skill farmers. The evidence presented in this paper does not fulfill the reformers' expectations in this regard. I have been unable to detect any rise in agricultural efficiency due to PROCEDE. Despite not being able to reject the null hypothesis of no efficiency effect, this paper's analysis of administrative landholding data shows that PROCEDE did induce a rise in the volume of agricultural land transactions. Furthermore, the theoretical results in this paper offers an unseen avenue through which efficiency gains may have been achieved.

In an environment with heterogenous agents, two sectors, and a sector-specific factor of production, liberalization of the market for that factor results in ambiguous outcomes for efficiency in each sector. Therefore, the welfare implications of establishing secure property rights over agricultural land are ambiguous from the perspective of the agricultural sector when examined in isolation.

However, I have shown that aggregate efficiency increases once the factor market is liberalized, so any fall in efficiency in the agricultural sector would be more than offset by enhancement of efficiency in the non-agricultural sector. Assuming that the theoretical model is a reasonably close representation of the experience of Mexico while the reforms were underway, the null empirical result on agricultural efficiency may indicate a rise in efficiency in the rest of the economy. Investigation of this possibility using data on off-farm income sources is a starting point for future work on the efficiency consequences of land titling in Mexico.

The lack of a detectable effect of land titling on agricultural efficiency is a cautionary result for policymakers. Land titling may not deliver any boost to agricultural efficiency. The theoretical results in this paper suggest that efficiency improvements may appear in the off-farm sector, but any boost in efficiency in the rest of the economy will likely be diffuse. Furthermore, receipt of credit fell as a result of land titling in this case. Land titling is supposed to expand credit access by allowing land to be used as collateral, but it is unclear whether that has happened here.

Finally, Mexico's land titling program caused landownership to become more concentrated. On the one hand, policymakers may want to encourage more land concentration to take advantage of scale economies and to encourage more people to enter the high-value manufacturing and service sector. On the other hand, more unequal distribution of land may be worrisome in a country that already experiences high income inequality.

References

[1] Ackerberg, D., Benkard, C. L., Berry, S., & Pakes, A. (2007). Econometric Tools for Analyzing Market Outcomes, In *Handbook of Econometrics*, Ed. James Heckman.

[2] Besley, T. (1995). "Property Rights and Investment Incentives: Theory and Evidence from Ghana." *Journal of Political Economy,* 103(5):903–37.

[3] Borjas, G. (1987). Economic Theory and International Migration. *International Migration Review*, 23(3), 457–485.

[4] Cord, L., & Wodon, Q. (2001). Do Agricultural Programs in Mexico Alleviate Poverty? Evidence from the Ejido Sector. *Cuadernos de Economia*, 256 (August), 239–256.

[5] Deaton, A. (1985). Panel Data from Time Series of Cross Sections. *Journal of Econometrics*, 30, 109–126.

[6] Deininger, K., Savastano, S., & Carletto, C. (2012). Land Fragmentation, Cropland Abandonment, and Land Market Operation in Albania. *World Development*, 40(10), 2108–2122.

[7] Diewert, W. E. (1999). Axiomatic and economic approaches to international comparisons. In *International and Interarea Comparisons of Income, Output, and Prices* (Vol. I), Ed. Alan Heston and Robert E. Lipsey.

[8] Feder, G. (1988). *Land Policies and Farm Productivity in Thailand*. Johns Hopkins University Press.

[9] de Ita, Ana (2006), Land Concentration in Mexico after PROCEDE. In *Promised Land: Competing Visions of Agrarian Reform*, Ed. Rosset, P.

[10] de Janvry, A., Emerick, K., Gonzalez-Navarro, M., & Sadoulet, E. (2015). Delinking Land Rights from Land Use: Certification and Migration in Mexico. *American Economic Review*, 105(10), 3125–3149.

[11] de Soto, H. (2000). *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. Basic Books.

[12] Deschênes, O. and M. Greenstone. (2007). The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather, *American Economic Review*. Vol. 97, No. 1, 354-385.

[13] Feder, G. and D. Feeny. (1991). Land Tenure and Property Rights: Theory and Implications for Development Policy. *The World Bank Economic Review*, Vol. 5, No. 1, 135-153.

[14] French, E., & Taber, C. (2011). Chapter 6 – Identification of Models of the Labor Market. *Handbook of Labor Economics*, (Vol. 4). Elsevier Inc.

[15] Galiani, S., & Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics*, 94(9–10), 700–729.

[16] Greene, W. H. (1981). On the Asymptotic Bias of Ordinary Least Square. *Econometrica*, 49(2), 505–513.

[17] Heath, J. R. (1992). Evaluating the Impact of Mexico's Land Reform on Agricultural Productivity. *World Development*, 20(5), 695–711.

[18] Heckman, J. J., & Honoré, B. E. (1990). The Empirical Content of the Roy Model. *Econometrica*, 58(5), 1121–1149.

[19] Holden, S. T., Deininger, K., & Ghebru, H. (2009). Impacts of low-cost land certification on investment and productivity. *American Journal of Agricultural Economics*, 91(2), 359–373.

[20] Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3), 346–363.

[21] Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics*, CXXIV(November), 1329–1372.

[22] Jin, S., & Deininger, K. (2009). Land rental markets in the process of rural structural transformation: Productivity and equity impacts from China. *Journal of Comparative Economics*, 37(4), 629–646.

[23] Johnson, N. L. (2001). Tierra y Libertad: Will Tenure Reform Improve Productivity in Mexico's Ejido Agriculture? *Economic Development and Cultural Change*, 49(2), 291–309.

[24] Kalirajan, K., Obwona, M., & Zhao, S. (1996). A decomposition of total factor productivity growth: the case of Chinese agricultural growth before and after reforms. *American Journal of Agricultural Economics,* 78(May), 331–338.

[25] Kumbhakar, S. (2001). Estimation of profit functions when profit is not maximum. *American Journal of Agricultural Economics*, 83(1), 1–19.

[26] Kumbhakar, S. & Lovell, C. (2000). *Stochastic Frontier Analysis.* Cambridge University Press.

[27] Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labor. The Manchester School 22.2: 139-191.

[28] Marschak, J., & Andrews, W. H. (1944). Random Simultaneous Equations and the Theory of Production. *Econometrica*, 12(3/4), 143–205.

[29] McCaa, Robert (2001). Missing millions: the human cost of the Mexican Revolution. *Mexican Studies* 19(2).

[30] Procuraduría Agraria (1998). Los tratos agrarios en ejidos certificados. Mexico DF.

[31] Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3(2), 135–146. http://doi.org/10.2307/2662082

[32] Ruane, A.C., R. Goldberg, and J. Chryssanthacopoulos, 2015. AgMIP climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation, *Agricultural and Forest Meteorology*, 200, 233-248.

[33] Sadoulet, E., Janvry, A. De, & Davis, B. (2001). Cash Transfer Programs with Income Multipliers: PROCAMPO in Mexico, International Food Policy Research Institute Food Consumption and Nutrition Division Discussion Paper 99.

[34] Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, 326–365.

[35] USDA (2013). PROCAMPO 2013 Subsidy Program Changes. Global Agricultural Information Network Report No. MX3012.

[36] Wineman, A. and L. S. Liverpool-Tasie. (2016). Land Markets and Equity of Land Distribution in Northwestern Tanzania. Selected Paper for the 2016 Annual Meeting of the Agricultural & Applied Economics Association.

[37] World Bank. (2001). Mexico Land Policy: A Decade after the Ejido Reform. Colombia, Mexico, and Venezuela Country Management Unit, Environmentally and Socially Sustainable Development Sector Management Unit, Latin America and the Caribbean Region, (Report 22187-ME).

APPENDIX: Theory

A. Effect of land market liberalization on agricultural efficiency

When land rights are granted, what happens to the mean of ϵ_F among the agents actually farming? I will show that mean agricultural efficiency rises in the case of positive selection or refugee selection and falls in the case of negative selection.

Denote $\left\{\epsilon_F^R, \epsilon_F^N\right\}$, the set of the skills of the agents with rights and those with no rights, as ϵ_F. We will want to consider the mean of the two groups together.

$E\left[\epsilon_F|\mathring{C} = F\right]$ denotes the expectation of skill levels in farming given that the agent has chosen farming before land rights have been granted. It is simply the weighted sum of the expectations of farming skill among those agents who have a more favorable outcome in farming and who have the capability of farming:

$E\left[\epsilon_F|\mathring{C} = F\right] = \mathring{P}^R \cdot E\left[\epsilon_F^R|\mathring{Y}_F^R > \mathring{Y}_B^R\right] + \mathring{P}^N \cdot E\left[\epsilon_F^N|\mathring{Y}_F^N > \mathring{Y}_B^N\right]$

where

$\mathring{P}^R$ is the proportion of farmers who originally had rights to land. This is equal to one in the first period.

$\mathring{P}^N$ is the proportion of farmers who originally had no rights to land. This is equal to zero in the first period.

Other variables are defined in the main text.

Simplification yields

$E\left[\epsilon_F|\mathring{C} = F\right] = E\left[\epsilon_F^R|\mathring{Y}_F^R > \mathring{Y}_B^R\right]$

Now I want to show how the type of selection determines the direction of the following inequality:

$E\left[\epsilon_F|\mathring{C} = F\right] \lesseqgtr E\left[\epsilon_F|\overset{*}{C} = F\right] = \overset{*}{P}^R \cdot E\left[\epsilon_F^R|\overset{*}{Y}_F^R > \overset{*}{Y}_B^R\right] + \overset{*}{P}^N \cdot E\left[\epsilon_F^N|\overset{*}{Y}_F^N > \overset{*}{Y}_B^N\right]$

where $E\left[\epsilon_F|\overset{*}{C} = F\right]$ is the expectation of skill levels in farming given that the agent has chosen farming after land rights have been granted.

In particular, I will show

$E\left[\epsilon_F|\mathring{C} = F\right] < E\left[\epsilon_F|\overset{*}{C} = F\right]$ in the case of positive selection or refugee selection and

$E\left[\epsilon_F|\mathring{C} = F\right] > E\left[\epsilon_F|\overset{*}{C} = F\right]$ in the case of negative selection.

To show $E\left[\epsilon_F|\mathring{C} = F\right] < E\left[\epsilon_F|\overset{*}{C} = F\right]$ it is sufficient to show that both $E\left[\epsilon_F^R|\mathring{Y}_F^R > \mathring{Y}_B^R\right] < E\left[\epsilon_F^R|\overset{*}{Y}_F^R > \overset{*}{Y}_B^R\right]$ and $E\left[\epsilon_F^R|\mathring{Y}_F^R > \mathring{Y}_B^R\right] < E\left[\epsilon_F^N|\overset{*}{Y}_F^N > \overset{*}{Y}_B^N\right]$.

Likewise, to have $E\left[\epsilon_F|\mathring{C} = F\right] < E\left[\epsilon_F|\overset{*}{C} = F\right]$ we merely need $E\left[\epsilon_F^R|\mathring{Y}_F^R > \mathring{Y}_B^R\right] > E\left[\epsilon_F^R|\overset{*}{Y}_F^R > \overset{*}{Y}_B^R\right]$

and $E\left[\epsilon_F^R|\mathring{Y}_F^R > \mathring{Y}_B^R\right] > E\left[\epsilon_F^N|\overset{*}{Y}_F^N > \overset{*}{Y}_B^N\right]$.

Recall these definitions:

$\overset{*}{Y}_F^R = \mu_F + \epsilon_F^R$

$\overset{*}{Y}_B^R = \mu_B + L + \epsilon_B^R$

Using these definitions we have:

$\overset{*}{Y}_F^R > \overset{*}{Y}_B^R$

$\mu_F + \epsilon_F^R > \mu_B + L + \epsilon_B^R$

$\mu_F - L + \epsilon_F^R > \mu_B + \epsilon_B^R$

Now examine the "no rights" potential incomes:

$\overset{*}{Y}_F^N = \mu_F - L + \epsilon_F^N$

$\overset{*}{Y}_B^N = \mu_B + \epsilon_B^N$

Therefore

$\overset{*}{Y}_F^N > \overset{*}{Y}_B^N \iff \mu_F - L + \epsilon_F^N > \mu_B + \epsilon_B^N$

Since $\begin{bmatrix} \epsilon_F^R & \epsilon_B^R \end{bmatrix}$ and $\begin{bmatrix} \epsilon_F^N & \epsilon_B^N \end{bmatrix}$ are identically distributed, we come to the following useful equality:

$$E\left[\epsilon_F^R|\overset{*}{Y}_F^R > \overset{*}{Y}_B^R\right] = E\left[\epsilon_F^N|\overset{*}{Y}_F^N > \overset{*}{Y}_B^N\right]$$

Hence we need only examine $E\left[\epsilon_F^R|\mathring{Y}_F^R > \mathring{Y}_B^R\right]$ and $E\left[\epsilon_F^R|\overset{*}{Y}_F^R > \overset{*}{Y}_B^R\right]$ to compare $E\left[\epsilon_F|\mathring{C} = F\right]$ and $E\left[\epsilon_F|\overset{*}{C} = F\right]$

In order to investigate only $E\left[\epsilon_F^R|\mathring{Y}_F^R > \mathring{Y}_B^R\right]$, change and simplify notation:

$Y_F = \mu_F + \epsilon_F$

$Y_B = \tilde{\mu}_B + \epsilon_B$

Let $\tilde{\mu}_B = \mu_B + L$

This change will allow the use of results that employ derivatives. The object of interest is:

$$\frac{\partial}{\partial \tilde{\mu}_B} E\left[\epsilon_F|Y_F > Y_B\right]$$

If $E\left[\epsilon_F^R|\mathring{Y}_F^R > \mathring{Y}_B^R\right] < E\left[\epsilon_F^R|\overset{*}{Y}_F^R > \overset{*}{Y}_B^R\right]$, then $\dfrac{\partial}{\partial \tilde{\mu}_B} E\left[\epsilon_F|Y_F > Y_B\right] > 0$, and vice versa.

We want to know how $E\left[\epsilon_F | Y_F > Y_B\right]$ responds when L rises (changes from zero to a positive value). To answer this question I will incorporate results from Heckman & Honoré (1990). First note that $\epsilon_F - \epsilon_B$ is distributed normal since ϵ_F and ϵ_B are distributed jointly normal. Because $\epsilon_F - \epsilon_B$ is normally distributed, its distribution is strictly log concave, which will allow the use of the following results.

Proposition 1 of Heckman & Honoré (1990) indicates that $0 < \frac{\partial}{\partial d} E\left[\epsilon_F - \epsilon_B | \epsilon_F - \epsilon_B > d\right] < 1$ for any d.

Equation 19 of Heckman & Honoré (1990) shows that

$$\frac{\partial}{\partial \tilde{\mu}_B} E\left[\epsilon_F | Y_F > Y_B\right] = a_F \left. \frac{\partial}{\partial d} E\left[\epsilon_F - \epsilon_B | \epsilon_F - \epsilon_B > d\right] \right|_{d=-\mu_F+\tilde{\mu}_B}$$

where $a_F = \dfrac{\sigma_F^2 - \sigma_{BF}}{\sigma_F^2 + \sigma_B^2 - 2\sigma_{BF}}$

Since $\left. \frac{\partial}{\partial d} E\left[\epsilon_F - \epsilon_B | \epsilon_F - \epsilon_B > d\right] \right|_{d=-\mu_F+\tilde{\mu}_B}$ is always positive, the sign of $\frac{\partial}{\partial \tilde{\mu}_B} E\left[\epsilon_F | Y_F > Y_B\right]$ is determined by the sign of a_F.

Let positive selection prevail. Then $a_F > 1$, so $\frac{\partial}{\partial \tilde{\mu}_B} E\left[\epsilon_F | Y_F > Y_B\right] > 0$ in this case.

Let refugee selection prevail. Then $0 < a_F < 1$, so $\frac{\partial}{\partial \tilde{\mu}_B} E\left[\epsilon_F | Y_F > Y_B\right] > 0$ in this case.

Let negative selection prevail. Then $a_F < 0$, so $\frac{\partial}{\partial \tilde{\mu}_B} E\left[\epsilon_F | Y_F > Y_B\right] < 0$ in this case.

B. Solution for the price of the unit of land

Next I will calculate the price L of a unit of land. I will use the solution in the next section. Obtaining the solution also demonstrates that a market equilibrium exists and is unique.

We wish to set the price L of a unit of land such that the market clears. In order for the market to clear, the number of agents that farm in the pre- and post-PROCEDE environment must be equal.

Assume that both the rights group and the no rights group have a unit mass.

Let $\overset{\circ}{P}{}_F^R$ be the proportion of the rights group that farms in the pre-PROCEDE period. Let $\overset{*}{P}{}_F^R$ and $\overset{*}{P}{}_F^N$ be the proportion of the rights group and no rights group, respectively, that farms in the post-PROCEDE period. Then the market clearing condition is:

$\overset{\circ}{P}{}_F^R = \overset{*}{P}{}_F^R + \overset{*}{P}{}_F^N$

Let $\sigma \equiv \sqrt{\sigma_F^2 + \sigma_B^2 - 2\sigma_{BF}}$. This is the standard deviation of $\epsilon_F - \epsilon_B$.

In the course of proving their Theorem 2, Heckman & Honoré (1990) show that

$$\overset{\circ}{P}{}_F^R = \Phi\left(\frac{E\left[\overset{\circ}{Y}{}_F^R\right] - E\left[\overset{\circ}{Y}{}_B^R\right]}{\sigma}\right) ; \overset{*}{P}{}_F^R = \Phi\left(\frac{E\left[\overset{*}{Y}{}_F^R\right] - E\left[\overset{*}{Y}{}_B^R\right]}{\sigma}\right) ; \overset{*}{P}{}_F^N = \Phi\left(\frac{E\left[\overset{*}{Y}{}_F^N\right] - E\left[\overset{*}{Y}{}_B^N\right]}{\sigma}\right)$$

Assuming $\mu_F = \mu_B$ will allow us to obtain a simple closed-form expression for L. With this assumption we have:

$$\overset{\circ}{P}{}_F^R = \Phi(0) ; \overset{*}{P}{}_F^R = \Phi\left(\frac{-L}{\sigma}\right) ; \overset{*}{P}{}_F^N = \Phi\left(\frac{-L}{\sigma}\right)$$

Now the market-clearing condition becomes

$0.5 = 2 \cdot \Phi\left(\frac{-L}{\sigma}\right)$

$0.25 = \Phi\left(\frac{-L}{\sigma}\right)$

Inverting the normal CDF yields

$-L/\sigma \approx -0.674$

$L \approx 0.674 \cdot \sqrt{\sigma_F^2 + \sigma_B^2 - 2\sigma_{BF}}$

L is positive except when both $\rho_{BF} = 1$ and $\sigma_F^2 = \sigma_B^2$, in which case $L = 0$.

The equality $Y_F = Y_B$ holds for all agents when $\rho_{BF} = 1$ and $\sigma_F^2 = \sigma_B^2$, which illustrates how it is a pathological case.

C. Effect of land market liberalization on economy-wide efficiency

I will now show that liberalization of the land market leads to greater efficiency in the economy as a whole regardless of how mean efficiency in each sector reacts.

Let $\overset{\square}{P}{}_C^k$ denote the mass of agents in period $\square \in \{\circ, *\}$ with initial land rights $k \in \{R, N\}$ who choose sector $C \in \{F, B\}$, Again assume that both the rights group and the no rights group have a unit mass.

Let P_F be the mass of all agents who farm and P_B the mass of all agents who build. These will take the same values in both periods since the number of land units is fixed and the land market clears. Hence

$P_F = \mathring{P}_F^R = \overset{*}{P}_F^R + \overset{*}{P}_F^N$

$P_B = \mathring{P}_B^R + \mathring{P}_B^N = \overset{*}{P}_B^R + \overset{*}{P}_B^N$

Assume $\mu_F = \mu_B$ since this condition will facilitate obtaining a closed-form solution.

$\overset{*}{P}_F^R = \overset{*}{P}_F^N = \Phi\left(\frac{-L}{\sigma}\right) = 0.25$ as shown in Section B.

$\mathring{P}_B^N = 1$ because all agents in the no rights group only have the option to build in the pre-PROCEDE period.

Therefore $\mathring{P}_F^R = \mathring{P}_B^R = 0.5$ and $\overset{*}{P}_B^R = \overset{*}{P}_B^N = 0.75$.

Since the no rights group does not have a choice as to their occupation in the initial period, the mean skill in the building sector for them is just the mean of ϵ_B^N:

$E\left[\epsilon_B^N | \mathring{Y}_B^R\right] = E\left[\epsilon_B^N\right] = 0$

Again let $\sigma \equiv \sqrt{\sigma_F^2 + \sigma_B^2 - 2\sigma_{BF}}$

Combining equation 12 and result R-1 of Heckman & Honore (1990) yields

$$E\left[\epsilon_F^R | \mathring{Y}_F^R > \mathring{Y}_B^R\right] = \left(\sigma_F^2 - \sigma_{BF}\right)\frac{1}{\sqrt{2\pi}}\exp\left\{\left(\left(E\left[\mathring{Y}_F^R\right] - E\left[\mathring{Y}_B^R\right]\right)/\sigma\right)^2/2\right\}/\Phi\left(\left(E\left[\mathring{Y}_F^R\right] - E\left[\mathring{Y}_B^R\right]\right)/\sigma\right)$$

$$= \left(\sigma_F^2 - \sigma_{BF}\right)\frac{2}{\sqrt{2\pi}}$$

since $E\left[\mathring{Y}_F^R\right] - E\left[\mathring{Y}_B^R\right] = 0$

By a similar calculation,

$E\left[\epsilon_B^R | \mathring{Y}_F^R < \mathring{Y}_B^R\right] = \left(\sigma_B^2 - \sigma_{BF}\right)\frac{2}{\sqrt{2\pi}}$

Hence across the entire economy the sum of skills actually employed in the pre-PROCEDE period is:

$\mathring{P}_F^R \cdot E\left[\epsilon_F^R | \mathring{Y}_F^R > \mathring{Y}_B^R\right] + \mathring{P}_B^R \cdot E\left[\epsilon_B^R | \mathring{Y}_F^R < \mathring{Y}_B^R\right] + \mathring{P}_B^N \cdot E\left[\epsilon_B^N | \mathring{Y}_B^R\right]$

$$= 0.5 \cdot \left(\sigma_F^2 - \sigma_{BF}\right)\frac{2}{\sqrt{2\pi}} + 0.5 \cdot \left(\sigma_B^2 - \sigma_{BF}\right)\frac{2}{\sqrt{2\pi}} + 1 \cdot 0$$

$$= \frac{2}{\sqrt{2\pi}}\left(\sigma_F^2 + \sigma_B^2 - 2\sigma_{BF}\right)$$

Now calculate the components of the skills employed in the post-PROCEDE period:

$$E\left[\epsilon_F^R|\overset{*}{Y}_F^R>\overset{*}{Y}_B^R\right]=\left(\sigma_F^2-\sigma_{BF}\right)\frac{1}{\sqrt{2\pi}}\exp\left\{\left(\left(E\left[\overset{*}{Y}_F^R\right]-E\left[\overset{*}{Y}_B^R\right]\right)/\sigma\right)^2/2\right\}/\Phi\left(\left(E\left[\overset{*}{Y}_F^R\right]-E\left[\overset{*}{Y}_B^R\right]\right)/\sigma\right)$$

$$=\left(\sigma_F^2-\sigma_{BF}\right)\frac{1}{\sqrt{2\pi}}\exp\left\{\left(\frac{-L}{\sigma}\right)^2/2\right\}/\Phi\left(\frac{-L}{\sigma}\right)$$

$$\approx\left(\sigma_F^2-\sigma_{BF}\right)\frac{5.02}{\sqrt{2\pi}}$$

And $E\left[\epsilon_F^R|\overset{*}{Y}_F^R>\overset{*}{Y}_B^R\right]=E\left[\epsilon_F^N|\overset{*}{Y}_F^N>\overset{*}{Y}_B^N\right]$ as previously shown.

$$E\left[\epsilon_B^R|\overset{*}{Y}_F^R<\overset{*}{Y}_B^R\right]=\left(\sigma_B^2-\sigma_{BF}\right)\frac{1}{\sqrt{2\pi}}\exp\left\{\left(\left(E\left[\overset{*}{Y}_B^R\right]-E\left[\overset{*}{Y}_F^R\right]\right)/\sigma\right)^2/2\right\}/\Phi\left(\left(E\left[\overset{*}{Y}_B^R\right]-E\left[\overset{*}{Y}_F^R\right]\right)/\sigma\right)$$

$$=\left(\sigma_B^2-\sigma_{BF}\right)\frac{1}{\sqrt{2\pi}}\exp\left\{\left(\frac{L}{\sigma}\right)^2/2\right\}/\Phi\left(\frac{L}{\sigma}\right)$$

$$\approx\left(\sigma_B^2-\sigma_{BF}\right)\frac{1.67}{\sqrt{2\pi}}$$

Because $E\left[\overset{*}{Y}_B^N\right]-E\left[\overset{*}{Y}_F^N\right]=E\left[\overset{*}{Y}_B^R\right]-E\left[\overset{*}{Y}_F^R\right]$, we have $E\left[\epsilon_B^R|\overset{*}{Y}_F^R<\overset{*}{Y}_B^R\right]=E\left[\epsilon_B^N|\overset{*}{Y}_F^N<\overset{*}{Y}_B^N\right]$

Hence across the entire economy the sum of skills actually employed in the post-PROCEDE period is:
$\overset{*}{P}_F^R\cdot E\left[\epsilon_F^R|\overset{*}{Y}_F^R>\overset{*}{Y}_B^R\right]+\overset{*}{P}_F^N\cdot E\left[\epsilon_F^N|\overset{*}{Y}_F^N>\overset{*}{Y}_B^N\right]+\overset{*}{P}_B^R\cdot E\left[\epsilon_B^R|\overset{*}{Y}_F^R<\overset{*}{Y}_B^R\right]+\overset{*}{P}_B^N\cdot E\left[\epsilon_B^N|\overset{*}{Y}_F^N<\overset{*}{Y}_B^N\right]$

$$\approx 0.25\cdot\left(\sigma_F^2-\sigma_{BF}\right)\frac{5.02}{\sqrt{2\pi}}+0.25\cdot\left(\sigma_F^2-\sigma_{BF}\right)\frac{5.02}{\sqrt{2\pi}}+0.75\cdot\left(\sigma_B^2-\sigma_{BF}\right)\frac{1.67}{\sqrt{2\pi}}+0.75\cdot\left(\sigma_B^2-\sigma_{BF}\right)\frac{1.67}{\sqrt{2\pi}}$$

$$\approx\frac{2.5}{\sqrt{2\pi}}\left(\sigma_F^2+\sigma_B^2-2\sigma_{BF}\right)$$

The change in the sum of skills in the economy from the pre- to the post-PROCEDE period is therefore approximately

$$\Delta_S=\frac{2.5}{\sqrt{2\pi}}\left(\sigma_F^2+\sigma_B^2-2\sigma_{BF}\right)-\frac{2}{\sqrt{2\pi}}\left(\sigma_F^2+\sigma_B^2-2\sigma_{BF}\right)$$

$$=\frac{1}{2\sqrt{2\pi}}\left(\sigma_F^2+\sigma_B^2-2\sigma_{BF}\right)$$

$$= \frac{1}{2\sqrt{2\pi}} \left(\sigma_F^2 + \sigma_B^2 - 2\rho_{BF} \cdot \sigma_F \cdot \sigma_B\right)$$

Δ_S is positive except when both $\rho_{BF} = 1$ and $\sigma_F^2 = \sigma_B^2$, in which case it is zero.

What parameter values will result in higher economy-wide skill enhancement? First, Δ_S is clearly decreasing in ρ_{BF}.

To examine how it responds to an increase in σ_F (note the symmetry of σ_F and σ_B), compute the derivative:

$$\frac{\partial \Delta_S}{\partial \sigma_F} = \frac{1}{\sqrt{2\pi}} (\sigma_F - \rho_{BF} \cdot \sigma_B)$$

Thus, if ρ_{BF} is nonpositive, an increase in σ_F unambiguously increases Δ_S. If $\sigma_F > \sigma_B$, we have the same outcome. If $\sigma_F \leq \sigma_B$ and $\rho_{BF} > 0$, then the sign of $\partial\Delta_S/\partial\sigma_F$ is ambiguous, but if $\rho_{BF} \neq 1$ then at least one of $\partial\Delta_S/\partial\sigma_F$ or $\partial\Delta_S/\partial\sigma_B$ will be positive.