Direct Measurement of Efficiency Gains from Land Titling: PROCEDE's Effect upon the Productivity of Mexican Agriculture

May 25, 2016

Travis McArthur

Department of Agricultural & Applied Economics,

University of Wisconsin-Madison

Contact email: tdmcarthur@wisc.edu

Selected Paper prepared for presentation for the 2016 Agricultural & Applied Economics Association, Boston, MA, July 31-August 2

Direct Measurement of Efficiency Gains from Land Titling: PROCEDE's Effect upon the Productivity of Mexican Agriculture*

Travis McArthur†

May 25, 2016

Abstract

From 1993 to 2006 Mexico granted land titles to over 3.6 million farmers via a systematic program named PROCEDE. Prior to the program, selling and renting land in most of Mexico's agricultural sector was prohibited. I estimate the impact of land market liberalization in this context. In an environment with heterogeneity in agents' skill levels in agricultural and nonagricultural activities, enabling land transactions will shift the skill composition of the agricultural sector, altering agricultural productivity. Allowing rural dwellers to pursue their comparative advantage, as PROCEDE did, affects productivity in a theoretically ambiguous way. Under a Roy model framework of occupational choice, those who are best at farming do not necessarily farm because they may be even better at nonagricultural activities. The sign of the effect of PROCEDE depends on the dispersion in skill levels. Hence, empirical work is needed to investigate the question. I use the 1991 and 2007 Mexican agricultural censuses to construct a panel of productivity at the community level. PROCEDE was implemented in nearly 30,000 communities at different times, so under the assumption that there are no time-trending community-specific characteristics that are correlated with both productivity changes and the PROCEDE rollout schedule, the effect of PROCEDE on productivity is identified via a community- and time-fixed effects specification. In a preliminary estimate, I find that PROCEDE substantially raised maize productivity.

*Selected Paper prepared for presentation at the 2016 annual meeting of the Agricultural & Applied Economics Association in Boston, MA, July 31 - August 2.

†Department of Agricultural & Applied Economics, University of Wisconsin-Madison. Email: tdmcarthur@wisc.edu.
Thanks to Brad Barham, Jean-Paul Chavas, Gustavo Gordillo, Cornelia Ilin, Alexey Kalinin, Matt Klein, Dan Phaneuf, Rashesh Shrestha, Chris Taber, and Srinivasan Vasudevan for helpful comments.

1 Introduction

From 1993 to 2006, the government of Mexico implemented a massive program to grant individual land titles to over 3.6 million farmers. The program aimed to facilitate the exercise of the new rights to sell, buy, and rent land in special communities called ejidos. These communities were established in the course of land redistribution in the early 20th century, but recipients of land grants did not receive individual land titles and had to agree to severe restrictions on land transactions. These restrictions, including prohibitions on the sale and rent of land, were lifted in 1992.

The land titling program, called PROCEDE, formed part of Mexico's package of free-market reforms in the late 1980's and early 1990's. A principal goal of the program was to increase efficiency in the agricultural sector by enabling the best farmers to obtain land while freeing those with a comparative advantage in other activities to pursue other occupations. According to prior work, PROCEDE accounted for a substantial portion of out-migration from ejidos to urban areas and the United States (de Janvry et al. 2015). If migrants' comparative advantage in farming factored into their decisions to leave these communities, PROCEDE's effects on out-migration can be taken as indirect evidence for efficiency changes. Other studies have examined PROCEDE's impact upon the frequency of land transactions and credit access, which, again, may have led to efficiency changes, but no work has directly struck at the productivity question.

I directly measure this land titling program on agricultural technical efficiency. I hypothesize that the mechanism at play is the reallocation of land to villagers with a comparative advantage in farming. My identification strategy relies upon time- and village-fixed effects estimation of an agricultural production function via a two-period (1991 and 2007) aggregate panel.

The paper will first discuss the institutional context of the ejido reforms. Then I will review prior work on PROCEDE in particular and technical efficiency issues in general. Next comes the theoretical model. The model, which relies on Roy (1951), demonstrates that the effect of land titling on mean technical inefficiency is ambiguous. Then I will outline a data generating process that describes how inputs and differing efficiency levels combine to produce agricultural output. An estimating equation for this generating process, which is intended to measure the "reduced form" effect of PROCEDE on efficiency, follows. Finally, I will examine evidence of PROCEDE's effect on land transactions by examining a database of agricultural subsidy recipients.

2 Background

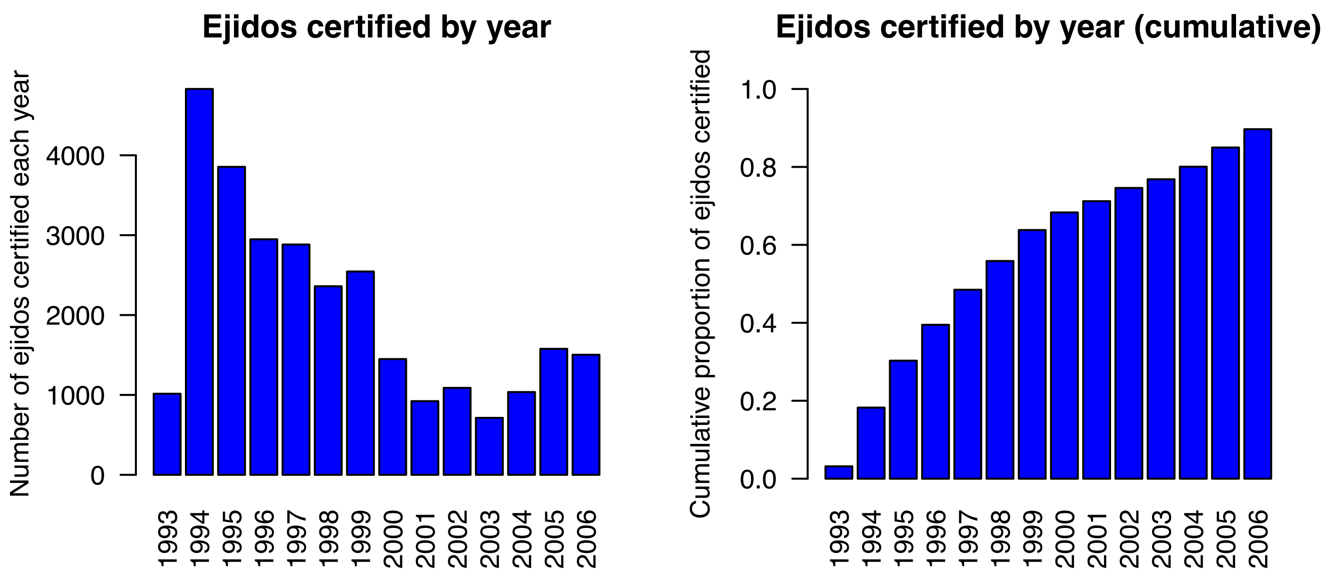
The Mexican Revolution, fought partially over inequality in landholdings, completely disrupted Mexican society and led to the deaths of 10 percent of the population (McCaa 2001). One of main outcomes of the Revolution was the creation of the legal framework for ejidos in 1917. The purpose of the ejido 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 were created over the entire 1920-1990 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).

The 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 receive 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, p. 151). The amendments were approved in late 1991 and became law in January 1992 (Johnson 2001). 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

Figure 1



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

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 (de Janvry et al. 2015). 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). A fixed effects strategy will remove any endogenous selection bias that results from preexisting observable or unobservable characteristics.

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 2011). 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) technical efficiency (or, in other words, agricultural output conditional on all inputs).

Work on the technical efficiency impacts of land tiling and land transactions include 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 have to resort to 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 and Klenow (2009) find Indian and Chinese manufacturers in 90th percentile of efficiency are four times more productive than those in 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 et al. (1996) find that Chinese agricultural total factor productivity rose by an annual average of eight 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. Liberation of the land market could accelerate this transition.

4 Theoretical model

Assume that farmers are heterogeneous in their skill level. 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 in off-farm activities. For simplicity, I will assume that individuals that exclusively engage in off-farm activities are synonymous with individuals who engage in off-farm activities and also farm, but who would prefer to cultivate more land. 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 wished 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 ejiditarios would have an absolute advantage in farming and another set an absolute advantage in off-farm activities. As more land shifts to ejiditarios 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 under the assumption that $f(\cdot)$ be 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. This will be a crucial representation for the goal of incorporating technical inefficiency into the standard form of the Roy model. This representation is

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

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 $-\frac{r}{1-r}$ in $\mathbf{w}$

It is easy to see from a simple algebraic standpoint that we need $r < 1$ because $r = 1$ results in a divide-by-zero issue 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 possible log earnings from building and let Y_{Fj} be the j'th individual's possible log earnings from farming. These are potential 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.

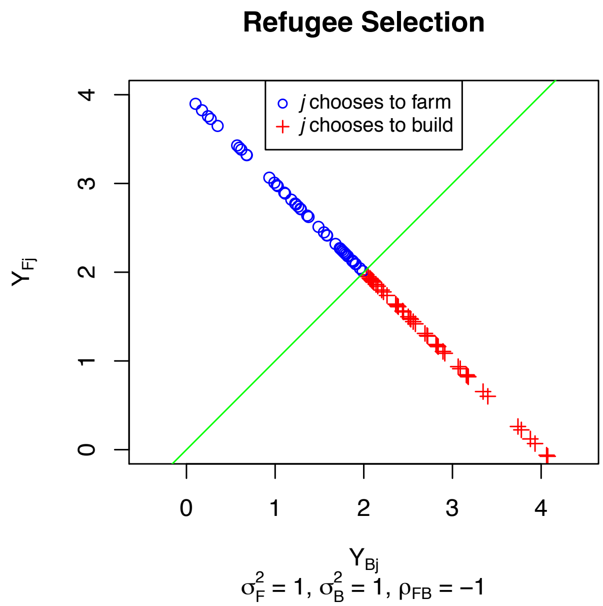
There are three possible scenarios depending on how agents are selected into farming. The first one, called refugee selection by the migration literature, occurs when the best farmers farm and the best builders build. The second, called positive selection, occurs when the best farmers farm, but the best 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 in which the best farmers are even better at building, so they choose to build; the not-so-great farmers choose to farm. 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 values of ϵ_{Bj} and ϵ_{Fj} . The plot character 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 out-of-equilibrium state where all agents

[1] A lecture by Chris Taber inspired these figures

are not free to choose occupations. After PROCEDE is implemented, an equilibrium with a liberated land market is established in which agents select into their preferred occupation.

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

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

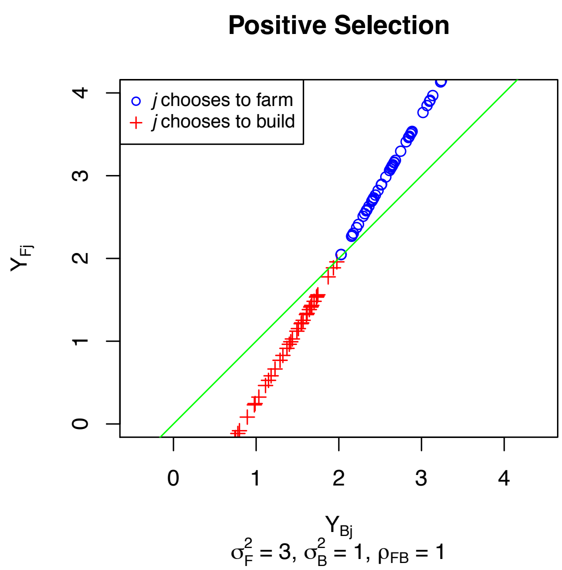
We can immediately return to the familiar environment of (2) by re-centering and re-scaling via these definitions:

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

This produces $Y_{Fj} = \mu_F + u_{Fj}$

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 $\begin{bmatrix} \epsilon^R_{Bj} & \epsilon^R_{Fj} \end{bmatrix}$ and $\begin{bmatrix} \epsilon^N_{Bj} & \epsilon^N_{Fj} \end{bmatrix}$ be independently and identically distributed as in (3). This is a somewhat unrealistic assumption. Those agents with usufruct rights inherited the rights from parents, and parents likely transferred farming skills to offspring.

Figure 3



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

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

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 C_j^{R*} is then determined by

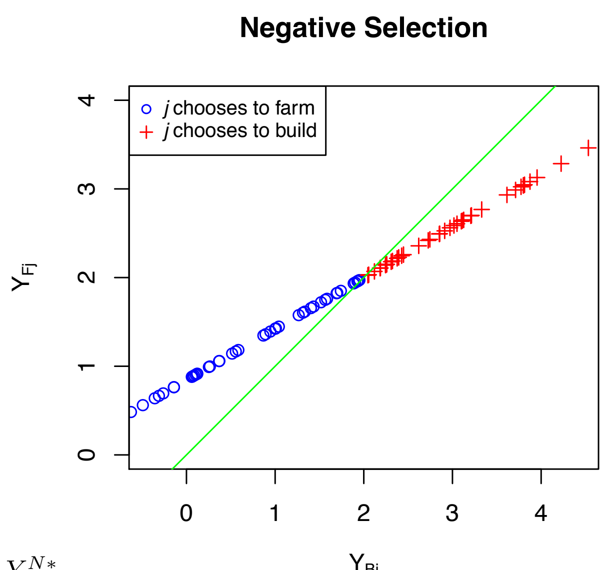
$$\begin{array}{l} Y_{Bj}^{R*} = \mu_B + \tilde{T} + \epsilon_{Bj}^R \\ Y_{Fj}^{R*} = \mu_F + \epsilon_{Fj}^R \end{array} \quad , \quad C_j^{R*} = \begin{cases} B & \text{if } Y_{Bj}^{R*} > Y_{Fj}^{R*} \\ F & \text{if } Y_{Bj}^{R*} \leq Y_{Fj}^{R*} \end{cases}$$

where $\tilde{T}$ is the market value of their land unit.

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 $\tilde{T} > 0$, clearly some of these agents will switch to $C_j^{R*} = B$. 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 $\tilde{T}$ is subtracted from Y_{Fj}^N in the post-PROCEDE decision rule:

Figure 4

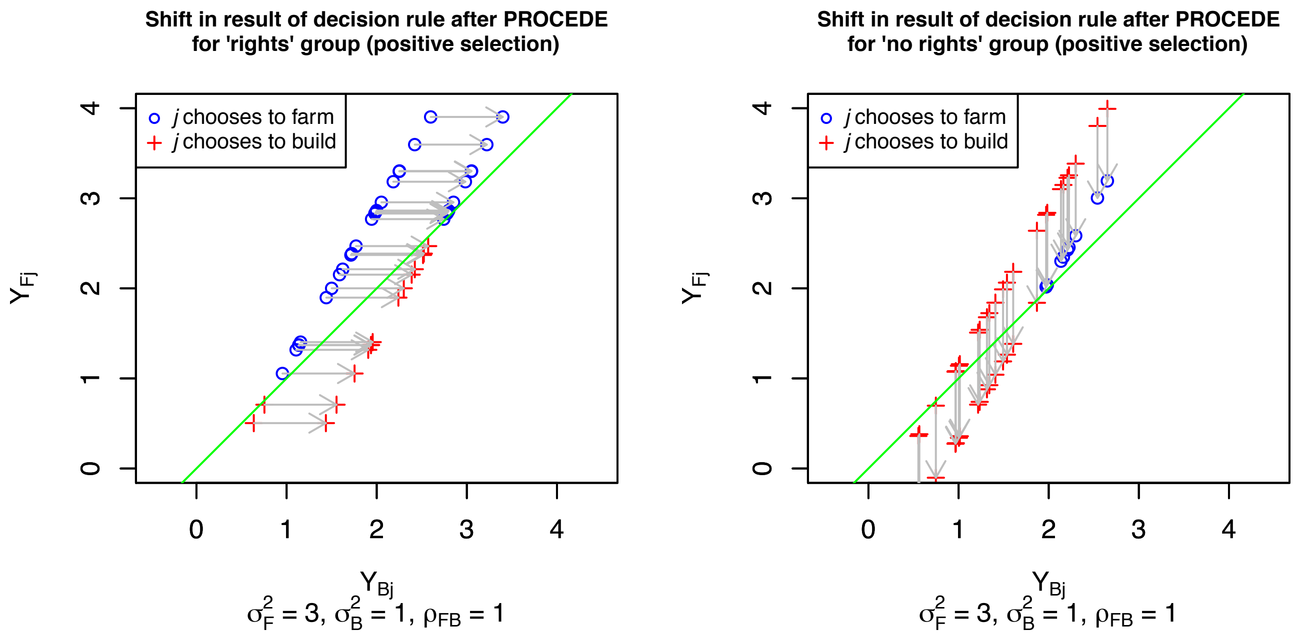


$\sigma_F^2 = 1$, $\sigma_B^2 = 3$, $\rho_{FB} = 1$

$$\begin{array}{l} Y_{Bj}^{N*} = \mu_B + \epsilon_{Bj}^N \\ Y_{Fj}^{N*} = \mu_F - \tilde{T} + \epsilon_{Fj}^N \end{array} \quad , \quad C_j^{N*} = \begin{cases} B & \text{if } Y_{Bj}^{N*} > Y_{Fj}^{N*} \\ F & \text{if } Y_{Bj}^{N*} \leq Y_{Fj}^{N*} \end{cases}$$

The shifts in the value $\mathring{Y}_{Bj}^R$ to Y_{Bj}^{R*} and $\mathring{Y}_{Fj}^N$ to Y_{Fj}^{N*} under conditions of positive selection are illustrated in Figure 5. The Y_{Bj} value of the "rights" group rises, while the Y_{Fj} value of the "no rights" group falls. 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 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 upward shift in this line indicates that technical efficiency has increased after PROCEDE under conditions of positive selection. Y_{Fj} is not exactly technical efficiency u_j, but Y_{Fj} is an affine transformation (given by (4)) of u_j, which

Figure 5



preserves ranking. Therefore, mean technical efficiency will have increased if the mean of $\{Y_{Fj} : C_j = F\}$ has increased. Figures 7 and 8 indicate that mean technical efficiency declines in the post-PROCEDE period when the error structure ensures negative selection.

It is clear from the figures that the rightholders' sale option unambiguously intensifies any selection (positive or negative) that exists. But the holders of usufruct rights do not sell their land into the ether. If we relax the i.i.d. assumption and allow the means, variances, or covariances of the ϵ^N_{Bj} and ϵ^N_{Fj} terms of the "no rights" group to substantially different from those of the "rights" group (ϵ^R_{Bj} and ϵ^R_{Fj}), the combined effect of the sale and purchase of land may be complicated. In any case, I have shown that under Roy model conditions, the expected effect of PROCEDE on productivity is ambiguous.

5 Data source & forming the aggregate panel

To estimate technical efficiency, I will 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

Figure 6

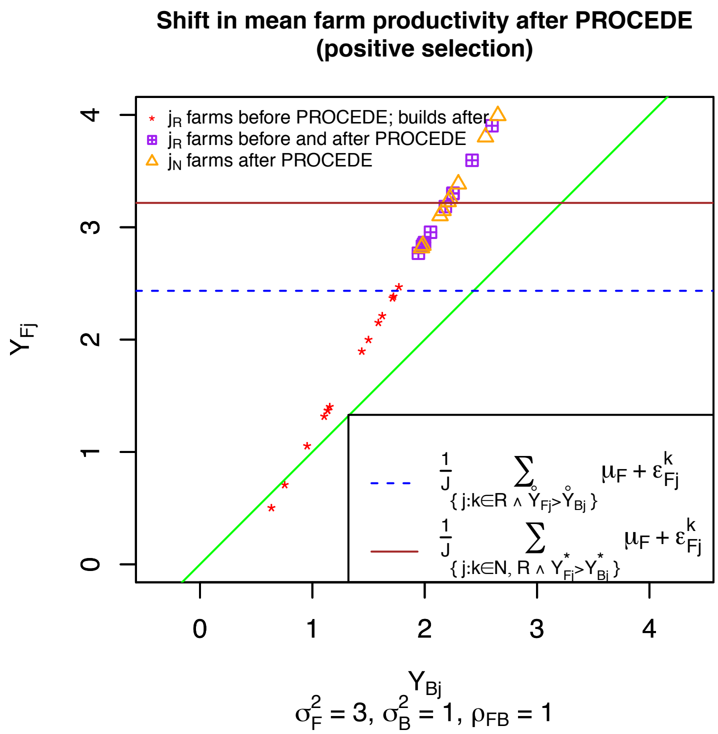


Figure 7

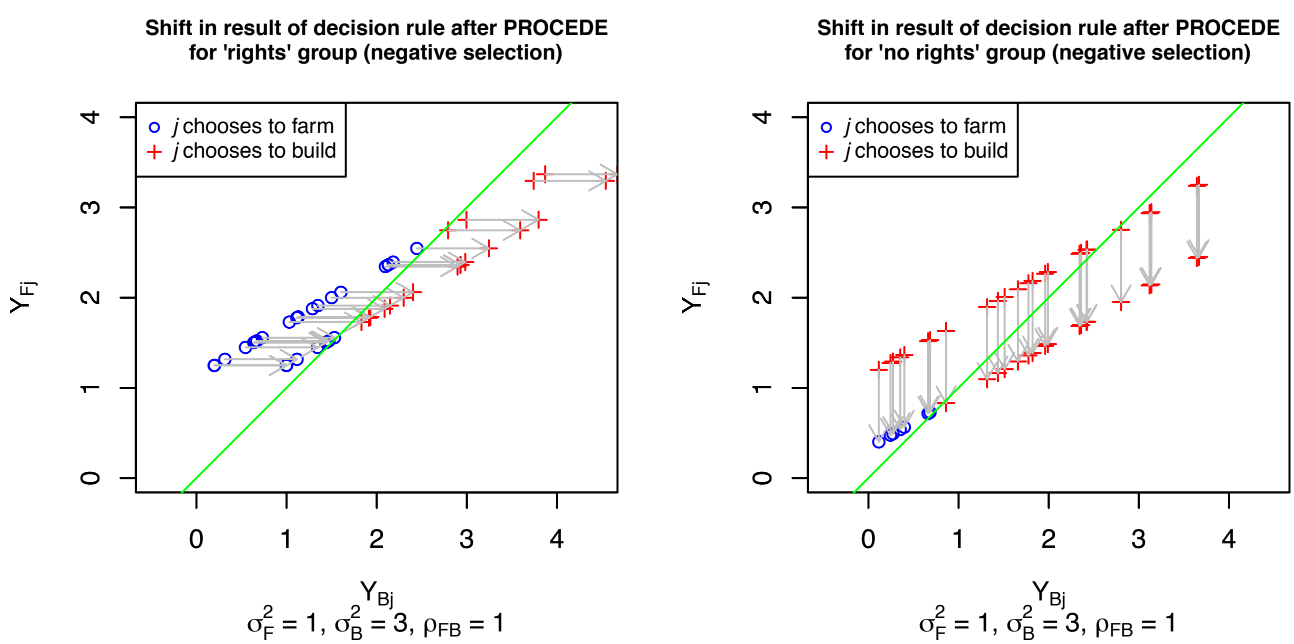
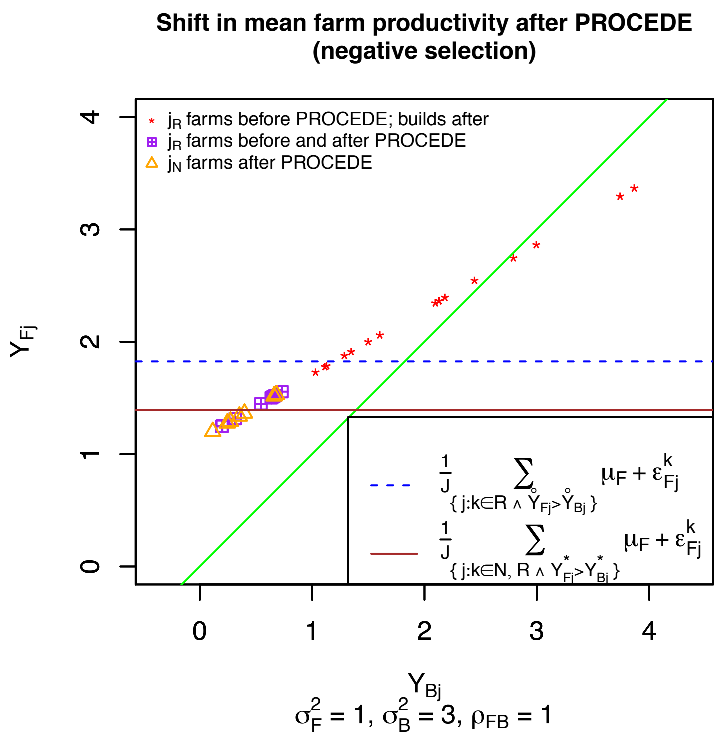


Figure 8


Shift in mean farm productivity after PROCEDE
(negative selection)
j_R farms before PROCEDE; builds after
j_R farms before and after PROCEDE
j_N farms after PROCEDE
$\frac{1}{J} \sum_{\{j:k \in R \wedge \mathring{Y}_{Fj} > \mathring{Y}_{Bj}\}} \mu_F + \varepsilon^k_{Fj}$
$\frac{1}{J} \sum_{\{j:k \in N, R \wedge Y^*_{Fj} > Y^*_{Bj}\}} \mu_F + \varepsilon^k_{Fj}$
0
1
2
3
4
Y_{Fj}
Y_{Bj}
$\sigma^2_F = 1, \sigma^2_B = 3, \rho_{FB} = 1$

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 their land that is irrigated. 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.

INEGI, the national statistics agency of Mexico, requires that access to the confidential microdata only occurs within their secure laboratory inside the agency's building itself. All data analysis dealing with the agricultural censuses occurred within this laboratory.

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 below.

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

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

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.

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 output of the j'th firm 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 firms,

Ω_{jvt} is a random variable.

Define $\vee_{vt}$ as the set of firms in the v'th ejido in year t. At this point, assume that firms 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 firm at the time input levels are chosen. It can be considered any firm- and time-specific shock that occurs too late in the production process for the firm 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 firm. Firms 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 firm and the t'th year. Firms 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 firm. Receipt of technical assistance, 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. Weather and natural disasters may comprise this term. 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

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

where

X^m_{jvt} is the level of the m'th input chosen by the j'th firm 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 output 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:

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

Now 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

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

Now assume that $\overline{\omega_{jvt}} = 0$. Time- and firm-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

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

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

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

$\overline{\epsilon_{jvt}}$ does not change because it is i.i.d. across firm, ejido, 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 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

$$\overline{y_{jvt}} = \hat{\beta}_0 + \sum_m \hat{\beta}_j \overline{X^m_{jvt}} + \sum_m \sum_n \hat{\beta_{ij}} \overline{X^m_{jvt} X^n_{jvt}} + \hat{\lambda}_t + \hat{\delta}_v + \hat{\gamma} \cdot PROCEDE_{vt} + \overline{\epsilon_{jvt}} \tag{8}$$

The time fixed effects λ_t absorbs the ζ_t productivity disturbance term.

Ejido fixed effects δ_v absorbs the u_v productivity disturbance error term and the mean of the $\overline{\psi_{jvt}}$ term.

If the variation in η_{vt} is due only to weather, I can match annual weather data to ejido coordinates and generate control variables that will eliminate possible endogeneity concerning the η_{vt} disturbance term. This is an object for further work.

Finally,

$$PROCEDE_{vt} = \begin{cases} 1 & \text{if } v \in \mathscr{P}_1 \text{ and } t = 2007 \\ 0 & \text{otherwise} \end{cases}$$

Hence, the $\hat{\gamma}$ estimate is expected to measure the effect of having PROCEDE in 2007 on average technical efficiency within an ejido.

8 Results

What follows are preliminary results. Due to the difficulty of geographic data matching, I aggregated data to the municipality level, which is a higher level of geography than ejidos. The PROCEDE status of a municipality was calculated as the proportion of ejidos with PROCEDE, weighted by their share of agricultural land area within the municipality. Analysis at the level of ejido is left for future work.

Table 1: Effect of PROCEDE on maize productivity

	Tons of maize harvested
PROCEDE.muni.prop	0.473*** (0.172)
Observations	4,070
R^2	0.791
Adjusted R^2	0.527
Residual Std. Error	1.973 (df = 1797)
Note:	*p<0.1; **p<0.05; ***p<0.01 Technological parameters omitted. Municipio- and year-fixed effects. Standard errors clustered at the municipio and year level via the Cameron-Gelbach-Miller method.

The results of the main estimating equation are displayed in Table 1. The estimated effect of PROCEDE implementation is a rise of 437 kg of maize production per farm once ejido and time fixed effects and all production inputs are accounted for. This is a large productivity gain, given that the average production per farm in 2007 was 8.3 metric tons. Average yield in 2007 was 4.9 tons per hectare, while average land under maize cultivation was 2.1 hectares.

A table with the estimated for all technological parameters is included in the appendix.

PROCEDE's impact on land transactions: 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). 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 ejidatarios 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 string probability matching, I was able to link these records to about 20,000 ejidos within a database on the timing of PROCEDE implementation. 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 we may infer that land transactions have likely taken place when we observe individuals decreasing their amount of land in PROCAMPO while simultaneously other individuals in the same ejido increase their PROCAMPO landholdings.

To formalize this notion of simultaneous rise and fall of individuals' landholdings, denote PROCAMPO landholdings for individual i in ejido v at time t as $L_{i,v,t}$. Then we can define the sum of all individuals' positive shifts in landholding from year to year in a given ejido as

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

The sum of negative shifts can be defined similarly:

$D^-_{vt} = \sum_{i \in v} L_{i,v,t} - L_{i,v,t-1} \cdot \mathbb{1}\{L_{i,v,t} < L_{i,v,t-1}\}$

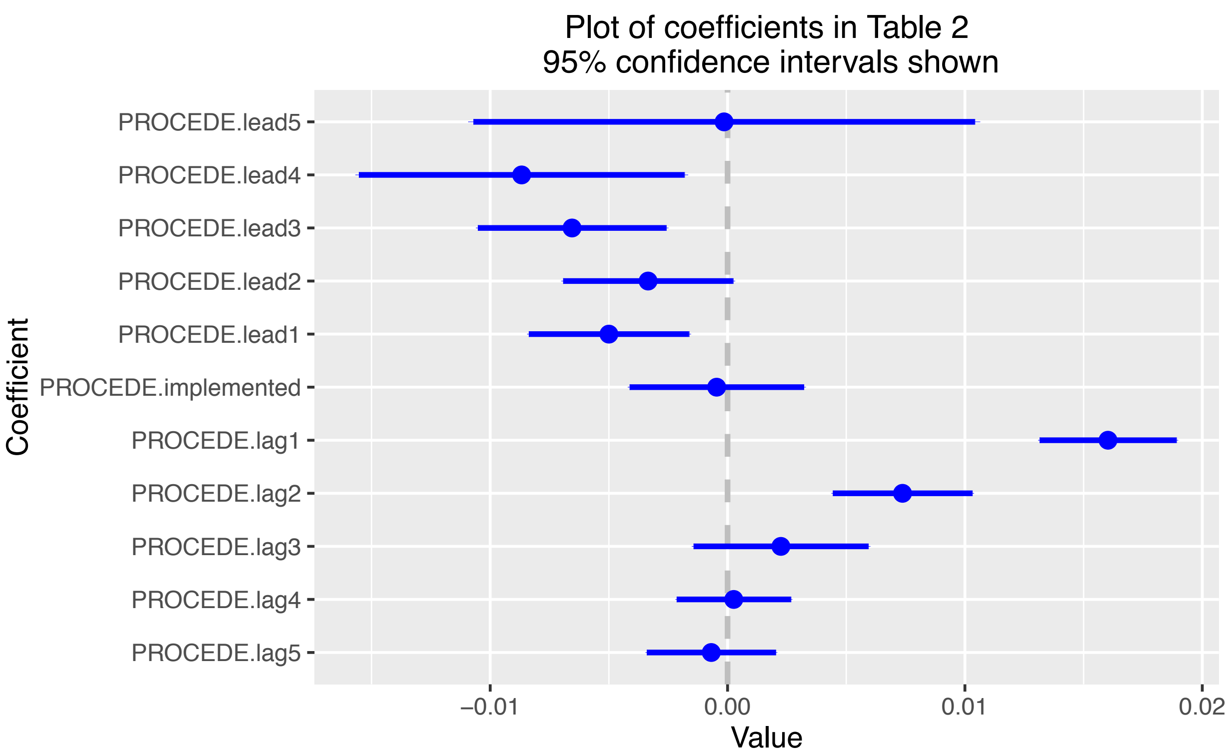
When D^+_{vt} and D^-_{vt} are both high in a given ejido and year, we may suspect that the volume of land transactions is also high. I defined a metric that is the square root of the product of D^+_{vt} and D^-_{vt} to formalize this notion. This metric is akin to an interaction term. I then regress this metric on dummies for lag and lead of PROCEDE implementation in each ejido, plus ejido and year fixed effects. Included as controls are D^+_{vt} and D^-_{vt} individually in order to difference out the individual effects of these metrics. I normalize these metrics by the total amount of land in the v'th ejido at time t.

Table 2

	$\frac{\sqrt{\text{hectares added to PROCAMPO} \times \text{hectares removed from PROCAMPO}}}{\text{hectares in PROCAMPO}}$
Hectares added/Total hectares	0.651*** (0.006)
Hectares removed/Total hectares	0.023*** (0.002)
PROCEDE.lag5	−0.001 (0.001)
PROCEDE.lag4	0.0003 (0.001)
PROCEDE.lag3	0.002 (0.002)
PROCEDE.lag2	0.007*** (0.001)
PROCEDE.lag1	0.016*** (0.001)
PROCEDE.implemented	−0.0005 (0.002)
PROCEDE.lead1	−0.005*** (0.002)
PROCEDE.lead2	−0.003* (0.002)
PROCEDE.lead3	−0.007*** (0.002)
PROCEDE.lead4	−0.009** (0.003)
PROCEDE.lead5	−0.0001 (0.005)
Observations	339,913
R^2	0.549
Adjusted R^2	0.519
Residual Std. Error	0.160 (df = 319081)
Note:	*p<0.1; **p<0.05; ***p<0.0
	Ejido and time fixed effects. Standard errors, in parentheses, clustered at the ejido leve

The results of this estimation, displayed in Table 2, suggest that there was a burst of transaction activity in the two years following PROCEDE implementation. Figure 9 aids interpretation of the trends. The transaction metric can be cautiously interpreted as the proportion of land in an ejido that changed hands in any given year. Thus, in the first year following PROCEDE implementation, the proportion of an ejido's

Figure 9


Plot of coefficients in Table 2
95% confidence intervals shown
Coefficient
PROCEDE.lead5
PROCEDE.lead4
PROCEDE.lead3
PROCEDE.lead2
PROCEDE.lead1
PROCEDE.implemented
PROCEDE.lag1
PROCEDE.lag2
PROCEDE.lag3
PROCEDE.lag4
PROCEDE.lag5
−0.01
0.00
0.01
0.02
Value

lands that changed hands rose 1.6 percent points above the baseline rate. The mean of this transaction metric is 7.6 percent, so the estimated effect of PROCEDE is substantial.

Since this transactions metric is admittedly unorthodox, I estimate models with different functional forms. These include the product of D^+_{vt} and D^-_{vt} without square root transformation and the minimum of D^+_{vt} and D^-_{vt}. I also use the rather discrete metrics of total number of recipients that had an increase/decrease in PROCAMPO landholdings and the total number of recipients that appeared or disappeared entirely in the database compared with the previous period. The results from these robustness checks, presented in the appendix, are broadly consistent with the specification presented in Table 2. These different flavors of the transaction metric are, naturally, strongly correlated, so they contribute only modestly to confidence in the results from the main specification.

9 Conclusion

Implementation of the PROCEDE land titling program substantially raised average agricultural productivity in Mexico. The mechanism for this shift cannot yet be pinned down conclusively. I have proposed one mechanism — sorting into agriculture due to heterogeneity of skills among agents. The apparent boost in land transactions following by PROCEDE, as indicated in an analysis of PROCAMPO recipient data, suggests that grant land titles indeed did facilitate the sale and rent of agricultural land. If the Roy model of occupation selection roughly describes conditions in Mexican ejidos, these dual empirical results would indicate that positive sorting into agriculture prevailed during this period.

References

[1] Ackerberg, D., Benkard, C. L., Berry, S., & Pakes, A. (2007). Econometric Tools for Analyzing Market Outcomes, 6(07). http://doi.org/10.1016/S1573-4412(07)06063-1

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

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

[4] Deaton, A. (1985). Panel data from time series of cross sections. *Journal of Econometrics*, 30, 109–126.

[5] Deininger, K., Savastano, S., & Carletto, C. (2012). Land Fragmentation, Cropland Abandonment, and Land Market Operation in Albania. *World Development*, 40(10), 2108–2122. http://doi.org/10.1016/j.worlddev.2012.05.010

[6] French, E., & Taber, C. (2011). Chapter 6 – Identification of Models of the Labor Market. HANDBOOK OF LABOR ECONOMICS, VOL 4A (Vol. 4). Elsevier Inc. http://doi.org/10.1016/S0169-7218(11)00412-6

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

[8] 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. http://doi.org/10.1257/aer.20130853

[9] Heath, J. R. (1992). Evaluating the impact of Mexico's land reform on agricultural productivity. *World Development*, 20(5), 695–711. http://doi.org/10.1016/0305-750X(92)90146-M

[10] Heckman, J. J., & Honoré, B. E. (1990). The Empirical Content of the Roy Model. *Econometrica*, 58(5), 1121–1149. http://doi.org/10.2307/2938303

[11] 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. http://doi.org/10.1111/j.1467-8276.2008.01241.x

[12] Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics*, CXXIV(November), 1329–1372. http://doi.org/10.1093/qje/qjs003

[13] 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. http://doi.org/10.1016/j.jce.2009.04.005

[14] 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. http://doi.org/10.1086/452503

[15] 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. Retrieved from http://www.jstor.org/stable/1243706

[16] Kumbhakar, S. (2001). Estimation of profit functions when profit is not maximum. *American Journal of Agricultural Economics*, 83(1), 1–19. Retrieved from http://ajae.oxfordjournals.org/content/83/1/1.short

[17] Kumbhakar, S. & Lovell, C. (2000). Stochastic Frontier Analysis.

[18] Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labor. The Economics of Underdevelopment. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9957.1954.tb00021.x/abstract

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

[20] McCaa, Robert (2001). Missing millions: the human cost of the Mexican Revolution. *Mexican Studies* 19.2

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

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

[23] Sadoulet, E., Janvry, A. De, & Davis, B. (2001). Cash Transfer Programs with Income Multipliers: PROCAMPO in Mexico, (99).

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

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

[26] 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

Code to perform the analyses in this paper is available at:

`github.com/tdmcarthur`

Technological parameters of the production function

Table 3 displays the results of the main estimation with all of the technological parameters Xj_x_Xk indicates the product of Xj and Xk. Table 4 displays the concordance between the inputs and their representation in the regression table. Many square terms were co-linear owing to the fact that input values were based on land area.

Table 3: Input concordance

X1	Total land in hectares
X2	Irrigated land
X3	Improved seeds
X4	Chemical fertilizer
X5	Organic fertilizer
X6	Pesticides and herbicides
X7	Animal traction
X8	Tractors
X9	Technical assistance
X10	Number of people working on the farm

Table 4: Effect of PROCEDE on maize productivity; all parameters

	Tons of maize harvested
Proportion of ejidos in municipality with PROCEDE	0.473*** (0.172)
X1	−0.184** (0.084)
X2	0.428 (0.327)
X3	0.804** (0.350)
X4	−0.042 (0.334)
X5	−0.255 (0.243)
X6	−0.074 (0.257)
X7	−0.361** (0.155)
X8	0.355*** (0.109)
X9	−2.030*** (0.487)
X10	−0.229*** (0.086)
X1_x_X1	0.0002 (0.0002)
X1_x_X2	−0.012 (0.010)
X1_x_X3	0.004 (0.005)

X1_x_X4	−0.001 (0.002)
X1_x_X5	−0.0003 (0.005)
X1_x_X6	−0.003 (0.008)
X1_x_X7	0.001* (0.0005)
X1_x_X8	−0.0001 (0.0002)
X1_x_X9	0.011 (0.010)
X1_x_X10	0.036 (0.027)
X2_x_X2	0.007 (0.007)
X2_x_X3	−0.002 (0.005)
X2_x_X4	−0.006 (0.005)
X2_x_X5	0.005* (0.003)
X2_x_X6	0.009*** (0.003)
X2_x_X7	0.005 (0.003)
X2_x_X8	0.005 (0.009)
X2_x_X9	0.001 (0.002)
X2_x_X10	−0.009 (0.013)
X3_x_X3	(0.000)
X3_x_X4	0.0004 (0.0005)
X3_x_X5	0.005 (0.006)
X3_x_X6	−0.0002 (0.003)
X3_x_X7	−0.006 (0.005)
X3_x_X8	−0.005 (0.005)
X3_x_X9	−0.026 (0.021)
X3_x_X10	0.024 (0.025)
X4_x_X4	(0.000)
X4_x_X5	0.002 (0.004)
X4_x_X6	0.002 (0.004)
X4_x_X7	−0.002 (0.002)
X4_x_X8	0.0004 (0.002)
X4_x_X9	−0.015** (0.006)
X4_x_X10	0.046 (0.033)
X5_x_X5	(0.000)
X5_x_X6	−0.004 (0.005)
X5_x_X7	−0.001 (0.005)
X5_x_X8	−0.001 (0.004)
X5_x_X9	−0.007 (0.005)
X5_x_X10	0.064** (0.030)
X6_x_X6	(0.000)
X6_x_X7	0.004 (0.004)
X6_x_X8	0.001 (0.005)
X6_x_X9	0.014*** (0.005)
X6_x_X10	−0.099*** (0.027)
X7_x_X7	(0.000)
X7_x_X8	−0.001** (0.001)
X7_x_X9	0.0005 (0.007)
X7_x_X10	−0.012 (0.022)
X8_x_X8	(0.000)
X8_x_X9	0.016 (0.023)
X8_x_X10	−0.017 (0.030)
X9_x_X9	(0.000)
X9_x_X10	0.025*** (0.009)

X10_x_X10	0.00002 (0.001)
Observations	4,070
R^2	0.791
Adjusted R^2	0.527
Residual Std. Error	1.973 (df = 1797)
Note:	*p<0.1; **p<0.05; ***p<0.01 Municipio- and year-fixed effects. Standard errors clustered at the municipio and year level via the Cameron-Gelbach-Miller method.

Alternative specifications for PROCEDE's effect on land transactions

Tables 5 and 6 display results for alternative measurements land transactions gleaned from the PROCAMPO beneficiary data.

Table 5: Alternative specifications for PROCEDE's effect on land transactions, I

	(1)	(2)	(3)	(4)
Hectares added/Total hectares	118.911***	0.532***		
	(3.508)	(0.004)		
Hectares removed/Total hectares	0.328***	0.001***		
	(0.103)	(0.0003)		
# who added land/Total recipients			30.762***	0.603***
			(0.739)	(0.005)
# who removed land/Total recipients			0.538***	0.039***
			(0.040)	(0.003)
PROCEDE.lag5	0.085	−0.0005	0.199***	−0.0001
	(0.224)	(0.001)	(0.062)	(0.001)
PROCEDE.lag4	−0.165	0.0003	0.063	0.0004
	(0.331)	(0.001)	(0.077)	(0.001)
PROCEDE.lag3	−0.711**	0.0001	−0.240***	0.005***
	(0.321)	(0.001)	(0.078)	(0.002)
PROCEDE.lag2	−0.095	0.004***	0.352***	0.013***
	(0.377)	(0.001)	(0.095)	(0.002)
PROCEDE.lag1	1.777***	0.012***	1.644***	0.027***
	(0.465)	(0.001)	(0.140)	(0.002)
PROCEDE.implemented	−0.523	−0.0002	−0.151	−0.002
	(0.407)	(0.001)	(0.095)	(0.002)
PROCEDE.lead1	−0.455	−0.003**	−0.230**	−0.006***
	(0.497)	(0.001)	(0.107)	(0.002)
PROCEDE.lead2	−1.159***	−0.002*	−0.252**	−0.005***
	(0.432)	(0.001)	(0.111)	(0.002)
PROCEDE.lead3	−1.014*	−0.003**	−0.403***	−0.008***
	(0.590)	(0.001)	(0.134)	(0.002)
PROCEDE.lead4	−2.388***	−0.008***	−0.653***	−0.012***
	(0.652)	(0.002)	(0.160)	(0.004)
PROCEDE.lead5	−1.983***	−0.001	−0.532***	−0.003
	(0.705)	(0.001)	(0.182)	(0.004)
Observations	339,913	339,913	339,913	339,913
R^2	0.217	0.608	0.227	0.564
Adjusted R^2	0.166	0.582	0.176	0.535

Note: *p<0.1; **p<0.05; ***p<0.01

Ejido and year fixed effects. Standard errors, in parentheses, clustered at the ejido level

LHS variable for (1) is $\frac{\text{hectares added to PROCAMPO} \times \text{hectares removed from PROCAMPO}}{\text{hectares in PROCAMPO}}$

LHS variable for (2) is $\frac{\min\{\text{hectares added to PROCAMPO}, \text{hectares removed from PROCAMPO}\}}{\text{hectares in PROCAMPO}}$

LHS variable for (3) is $\frac{\text{\# of recipients who added land} \times \text{\# of recipients who removed land}}{\text{\# of PROCAMPO recipients}}$

LHS variable for (4) is $\frac{\sqrt{\text{\# of recipients who added land} \times \text{\# of recipients who removed land}}}{\text{\# of PROCAMPO recipients}}$

Table 6: Alternative specifications for PROCEDE's effect on land transactions, II

	(1)	(2)	(3)	(4)
# who added land/Total recipients	0.480***			
	(0.003)			
# who removed land/Total recipients	0.005***			
	(0.0003)			
# who appeared/Total recipients		34.374***	0.632***	0.528***
		(0.904)	(0.005)	(0.004)
# who disappeared/Total recipients		0.355***	0.027***	0.003***
		(0.033)	(0.003)	(0.0002)
PROCEDE.lag5	−0.001	0.129**	−0.001	−0.0003
	(0.001)	(0.056)	(0.001)	(0.001)
PROCEDE.lag4	0.0001	0.038	−0.0003	0.0002
	(0.001)	(0.065)	(0.001)	(0.001)
PROCEDE.lag3	0.001	−0.170**	0.002*	0.0003
	(0.001)	(0.069)	(0.001)	(0.001)
PROCEDE.lag2	0.008***	−0.148**	0.002**	0.002***
	(0.001)	(0.075)	(0.001)	(0.001)
PROCEDE.lag1	0.020***	−0.045	0.008***	0.006***
	(0.001)	(0.091)	(0.001)	(0.001)
PROCEDE.implemented	−0.0005	−0.155*	−0.001	−0.00004
	(0.001)	(0.084)	(0.001)	(0.001)
PROCEDE.lead1	−0.003***	−0.214**	−0.005***	−0.002**
	(0.001)	(0.097)	(0.001)	(0.001)
PROCEDE.lead2	−0.002**	−0.279***	−0.006***	−0.003**
	(0.001)	(0.103)	(0.002)	(0.001)
PROCEDE.lead3	−0.003**	−0.421***	−0.007***	−0.002**
	(0.001)	(0.126)	(0.002)	(0.001)
PROCEDE.lead4	−0.007***	−0.759***	−0.012***	−0.008***
	(0.002)	(0.151)	(0.003)	(0.001)
PROCEDE.lead5	−0.003**	−0.583***	−0.002	−0.002*
	(0.001)	(0.158)	(0.004)	(0.001)
Observations	339,913	339,913	339,913	339,913
R^2	0.593	0.220	0.526	0.606
Adjusted R^2	0.567	0.169	0.495	0.580

Note: *p<0.1; **p<0.05; ***p<0.01

Ejido and year fixed effects. Standard errors, in parentheses, clustered at the ejido level

LHS for (1) is $\frac{\min\{\text{\# of recipients who added land}, \text{\# of recipients who removed land}\}}{\text{\# of PROCAMPO recipients}}$

LHS for (2) is $\frac{\text{\# of people who appeared in dataset} \times \text{\# of people who disappeared from dataset}}{\text{\# of PROCAMPO recipients}}$

LHS for (3) is $\frac{\sqrt{\text{\# of people who appeared in dataset} \times \text{\# of people who disappeared from dataset}}}{\text{\# of PROCAMPO recipients}}$

LHS for (4) is $\frac{\min\{\text{\# of people who appeared in dataset}, \text{\# of people who disappeared from dataset}\}}{\text{\# of PROCAMPO recipients}}$