ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF BUSINESS, ECONOMICS AND LAW
UNIVERSITY OF GOTHENBURG
207

Essays in Political Economy of Development

Michele Valsecchi

UNIVERSITY OF GOTHENBURG

ISBN 978-91-85169-69-6
ISSN 1651-4289 print
ISSN 1651-4297 online

Printed in Sweden,
Kompendiet 2012

A chi insiste,
persiste,
e poi, esausto, desiste

Contents

Paper 1:

Social Conflict, Fractionalization, and Polarization

Paper 2:

Land Property Rights and International Migration: Evidence from Mexico

Paper 3:

Local Elections and Corruption during Democratization: Evidence from Indonesia

Paper 4:

Resource Windfalls and Public Goods: Evidence from a Policy Reform

Acknowledgements

First of all, I wish to thank my supervisors, Ola Olsson and Måns Söderbom, for all inspiring discussions, ideas, feedback and guidance through the dissertation process. This thesis would not have been possible without the assistance of both of you.

I am also grateful to everybody at the Department of Economics for their comments during my seminars, their cheer-up jokes in the corridor (some of them do make good jokes actually) and the general positive atmosphere they promote. Thanks to Lisa, Kristina, Haile, Annika, Conny, Sven, Måns, Yonas for support (and forgive me if I forget some of you).

Clearly I am grateful to all the administrators I happened to ask questions and support: Eva-Lena, who helped me about everything; Jeanette, who helped with about everything; Åsa, who explained me my contract several times although I still don't fully get it. Of course I want to thank Debbie for her patience.

I gratefully acknowledge financial assistance from The Royal Society of Arts and Sciences in Göteborg (KVVS) and from Knut and Alice Wallenberg Foundation.

Finally, I want to thank my parents and my brother, for putting up with me when I came to visit them and ended up working from their place,[1] and, above all, Erika, for putting up with me basically all the time. I suspect that I might have been not easy to deal with..

Michele Valsecchi

October 2012

[1] I also want to thank my dogs, Trissy and Whisky, for supporting me unconditionally since the very first day.

Summary of the thesis

The thesis consists of four self-contained papers.

Paper 1:

Social Conflict, Fractionalization, and Polarization

We develop a conflict model linking conflict intensity to the distribution of the population over an arbitrary number of groups. We extend the pure contest version of the model by Esteban and Ray (1999: "Conflict and Distribution", Journal of Economic Theory, 87(2): 379-415) to include a mixed public-private good. We analyze how the level of dissipation changes as the population distribution and the share of publicness of the prize change. In contrast to Esteban and Ray (2011: "Linking Conflict to Inequality and Polarization", American Economic Review, 101(4): 1345–74), we do not assume that the probability of winning equals group size. First, we characterize how the global maximum varies with the degree of publicness of the prize. Second, we find that, in case of pure private goods, the conflict-distribution relationship resembles the fractionalization index. Finally, we find that smaller groups always contribute more and so the fractionalization index underestimates their weight. Indeed, we find that the fractionalization index underestimates the true level of conflict.

Paper 2:

Land Property Rights and International Migration: Evidence from Mexico

In this paper we ask whether there is a relationship between land property rights and international migration. In order to identify the impact of property rights, we consider a country-wide land certification program that took place in Mexico in the 1990s. Our identification strategy exploits the timing of the program and the heterogeneity in farmers' eligibility for the program. Comparing eligible and ineligible households, we find that the program increased the likelihood of having one or more members abroad by 12 percent. In terms of number of migrants, our coefficient estimates explain 31 percent of the 1994-1997 increase in migrants from ejido areas and 16-18 percent of the increase from the entire Mexico. We contribute to the current debate on the determinants of Mexican emigration (Hanson 2006, Hanson and McIntosh 2009, Hanson and McIntosh 2010). Consistent with our theoretical model, the impact is strongest for households without a land will.

Paper 3:

Local Elections and Corruption during Democratization: Evidence from Indonesia

In this paper we ask whether the direct election of the local government increases accountability and decreases corruption. In order to identify the causal effect of direct elections, we exploit the gradual introduction of local elections in Indonesia and a novel dataset of corruption events that covers all districts during the period 1998-2008. We find that direct elections increase the number of corruption crimes by about half the pre-election average. We also find that embezzlement practices dominate all other types of corruption activities.

Paper 4:

Resource Windfalls and Public Goods: Evidence from a Policy Reform

In this paper, we outline an empirical approach for understanding whether natural resource windfalls have a positive or negative impact on local governments' provision of public goods. The literature on the curse of natural resources suggests that resource windfalls might not necessarily lead to good economic outcomes and that rents might be squandered in corruption and rent seeking. In order to identify the impact of natural resources on local government behavior, we exploit a country-wide fiscal decentralization reform in Indonesia, providing producing provinces a direct share of resource revenues. Our identification strategy is to compare villages along the border of three producing provinces in Sumatra and Kalimantan before and after the legislative change. Detailed descriptive statistics on district government budgets confirm the goodness of the research design. Regression analysis on a wide range of public goods suggests that the revenue windfall had a positive impact on the prevalence of high schools and various other public goods. We find no evidence of a resource curse.

Paper 1

Social Conflict, Fractionalization, and Polarization

Michele Valsecchi*

University of Gothenburg

First version: April 2010

This version: September 2012

Abstract

We develop a conflict model linking conflict intensity to the distribution of the population over an arbitrary number of groups. We extend the pure contest version of the model by Esteban and Ray (1999: "Conflict and Distribution", Journal of Economic Theory, 87(2): 379-415) to include a mixed public-private good. We analyze how the level of dissipation changes as the population distribution and the share of publicness of the prize change. In contrast to Esteban and Ray (2011: "Linking Conflict to Inequality and Polarization", American Economic Review, 101(4): 1345–74), we do not assume that the probability of winning equals group size. First, we characterize how the global maximum varies with the degree of publicness of the prize. Second, we find that, in case of pure private goods, the conflict-distribution relationship resembles the fractionalization index. Finally, we find that smaller groups always contribute more and so the fractionalization index underestimates their weight. Indeed, we find that the fractionalization index underestimates the true level of conflict.

Key words: ethnic diversity, public-private goods, polarization, fractionalization.

JEL Classification codes: D72, D73, D74, H42.

*Email: michele.valsecchi@economics.gu.se. I am indebted to Ola Olsson and Mario Gilli for many useful discussions. I have benefited from comments by Francesco de Sinopoli, Masayuki Kudamatsu and Måns Söderbom. I am grateful to Debraj Ray for comments on an earlier draft. I am also grateful to seminar participants at the University of Gothenburg, the "Global costs of conflict" workshop 2010 at DIW Berlin, and the CSAE 2010 conference in Oxford. I ackowledge financial support from Adlerbertska stipendiefonden and Paul och Marie Berghaus donationsfond. All remaining errors are my own.

1 Introduction

This paper draws inspiration from the mixed findings on the effects of ethnic diversity on conflict and economic outcomes: ethnic fractionalization affects negatively economic performance (Easterly and Levine (1997)), while ethnic polarization does not (Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003)); ethnic polarization has a negative effect on civil war incidence, while ethnic fractionalization does not (Montalvo and Reynal-Querol (2005)).[1] The sensitiveness of these relationships to the index used to capture ethnic diversity may inform us as to the mechanisms through which they work. The research questions we tackle in this paper are: which sorts of distributions are associated with high levels of conflict? Does the conflict-distribution relationship resemble the fractionalization or polarization index? In order to answer these questions, we develop a behavioral model linking societal conflict to the distribution of a population across groups and also investigate how societal conflict changes as the population distribution changes.

We conceive societal conflict as a situation where in presence of weak institutions (absence of checks and balances, absence of elections, or inefficiency of elections to discipline politicians) and in absence of a well-defined and agreed-upon collective decision rule, individuals incur costs to capture their most preferred outcome. The concept encompasses both rent-seeking behavior, i.e., lobbying, and open conflict. We study a simple rent-seeking model with an arbitrary number of groups. The characteristic feature of this class of models is the diversion of resources from productive activities.

The model borrows largely from the pure contest version of the model by Esteban and Ray (1999) (henceforth ER1999), who investigate the relationship between conflict and distribution. Since the properties of their model resemble closely those of the polarization index, one way to answer our research questions is to extend it in a way that makes the properties

[1] Magnitude and significance of these relationships are, to a certain extent, sensitive to the source of data on ethnicity and conflict. Alesina and Ferrara (2005) and Blattman and Miguel (2010) discuss these and other related issues. See also Valsecchi (2010) for some sensitivity tests on these relationships. Campos and Kuzeyev (2007), Ahlerup and Olsson (2012), Spolaore and Wacziarg (2009), and Michalopoulos (2012) study the determinants of ethnic diversity. Fearon (2003), Caselli and Coleman (2006), and Esteban and Ray (2008) explain why ethnic diversity may be particularly salient.

of the model resemble the fractionalization index for some parameter values, and those of the discrete polarization index for some others. By doing so, the model should suggest which features drive the change in the properties and which ones do not matter. The main novelty with respect to their model is the specification of the prize. Within the winning group, part of the outcome is a public good and that is enjoyed in the same quantity by all group members, regardless of their number; another part is private, in the sense that it has to be shared among group members, which means that the per capita share shrinks with group size.

A recent work by Esteban and Ray (2011) extends the framework of ER1999 along similar lines.[2] They find that a monotone transformation of the equilibrium level of conflict is a function of the Gini coefficient, the fractionalization index, and the polarization index. However, in order to reach this result, they have to assume that one of the endogenous variables (winning probabilities) equals one of the parameters (group size). The main difference between their paper and ours is that we do not impose this assumption. In the spirit of ER1999, the link between the theory and the indexes in our paper is based on the comparison between the properties of the model (our comparative statics) and the properties of the indexes.

We settle the existence and uniqueness of the equilibrium (Proposition 1) and we turn to the properties of the model. Our first observation is that, in case of two groups, conflict is always maximized when the population is uniformly distributed across them (Proposition 2.1). Our second observation is that, over the set of uniform distributions, the number of groups maximizing conflict decreases with the degree of publicness of the prize (Proposition 2.2). This property already suggests that the modeling choice regarding the prize at stake has a bite.

We characterize more precisely the relationship between the model and the indexes when the prize is a pure private good. In this case the model shares most of the properties of the fractionalization index (Proposition 3). This reinforces the theoretical grounds for the use of this index in empirical applications. In light of existing empirical evidence and earlier findings by ER1999, it also suggests that the key difference between the mechanism through which ethnic diversity affects economic performance and the incidence of civil wars is the nature of

[2] We explain the differences between their model and ours in Section 3.

the prize at stake.

The model in the present paper is more flexible than the fractionalization index since it allows members of different groups to devote different contributions. Therefore we ask: Do all groups devote the same per capita contributions or do some contribute more than others? We find that members of smaller groups always contribute more (Proposition 4). Thus, the fractionalization index systematically underestimates the weight of smaller groups in the creation of conflict. Indeed, we find that, in a special case, the fractionalization index underestimates the true level of conflict (Proposition 5).

2 Diversity: measures and properties

The fractionalization index is the probability that any two randomly chosen individuals belong to different ethnic groups. Let the size of a generic group be denoted by n_i and the entire population be normalized to unity $\left(\sum_{i=1}^{G} n_i = 1\right)$. Then the fractionalization index is:[3]

$$F = \sum_{i=1}^{G} n_i (1 - n_i) = 1 - \sum_{i=1}^{G} n_i^2.$$

It has the following properties:

1. for a given number of groups G, F is maximized at the uniform population distribution over these groups;
2. over the set of uniform distributions, F increases with the number of groups;
3. the splitting of any group into two new groups increases F;
4. any transfer of population to a smaller group increases F.

Since the impact of a split (3) on the index depends neither on the size of the group that is split nor on the distribution of the other groups, the index is said to be *local*. Properties

[3] The index has two theoretical backgrounds: one is the Gini coefficient (the fractionalization index can be seen as its semplification), and the other is the Herfindal index (the fractionalization index is its complement).

3 and 4 imply that it is always possible to break down a transfer into a sequence of smaller transfers, all changing the index in the same direction. For this reason the index is said to be *monotonic*.

The discrete polarization index is a simplified version of the polarization index introduced by Esteban and Ray (1994).[4] The expression for its discrete version (Q) is:

$$Q = 4\sum_{i=1}^{G}\left[n_i^2\left(1-n_i\right)\right],$$

where n_i denotes the population share for group i and the population is normalized to unity: $\sum_{i=1}^{G} n_i = 1$.

It has the following properties:

1. for a given number of groups G, Q is maximized when the population is concentrated in two equally sized groups only (bimodal symmetric distribution);

2. over the set of uniform distributions, Q decreases with the number of groups, provided there are at least two groups to begin with;

3. the splitting of a group in two increases Q if and only if the initial group size was at least 2/3;

4. a transfer of population to a smaller group increases Q if both groups are larger than 1/3. If both groups are smaller than 1/3, the transfer decreases Q.

Since the impact of a split (3) on the index depends on the size of the non-splitting population, which is not directly associated with the change, the index is said to be *global* (Esteban and Ray 1994:829). Properties 3 and 4 imply that a population change cannot necessarily be broken down into a sequence of changes having the same effect on the index. For this reason the index is said to be *non-monotonic* (Esteban and Ray 1994:829).

[4] Essentially, Reynal-Querol and Montalvo (2002) and Montalvo and Reynal-Querol (2005) simplified the expression for the general index to exclude the use of ethnic distances, normalized the index to unity to make it easier to be interpreted, and chose a particular value of a polarization sensitiveness (see one of the paper for details). Note that the main purpose of the latter was to provide an alternative to the Gini coefficient in the field of inequality measurement and that the fractionalization index constitutes a semplification of the Gini coefficient itself.

Note that in case $G = 2$, both measures reach their maximum in correspondence of the uniform distribution ($n_1 = n_2 = 1/2$) and transfers from big to small groups increase both indexes.[5] The two indexes diverge more and more as the number of groups with positive population shares increases ($G \geq 3$), since Q maintains its maximum in correspondence of the bimodal distribution (population concentrated in any two groups with equal population shares $n_i = n_j = 1/2$), while the maximum for F becomes the uniform distribution over all groups.

3 The model

We provide a behavioral model linking conflict to the distribution of the population over a set of groups. We consider the pure contest version of the model by ER1999. Individuals belonging to different groups compete for the capture of a prize. We extend their model by specifying a mixed public-private prize. This feature introduces an additional channel through which group size determines the incentives of economic agents to contribute. Group size determines the per capita share of the private component: the bigger the group, the smaller the per capita share.[6] Whether this means introducing the Pareto-Olson argument into the model will be discussed later in the section.

Esteban and Ray (2011) also introduce a mixed public-private good in their 1999 framework, along with varying intra-group cohesion and inter-group distances. They find that the equilibrium level of conflict is a linear function of the Gini coefficient, the fractionalization and the polarization index. In order to reach this finding, they have to assume that one of the endogeneous variables (the probability of winning) equals one of the parameters of the model (group sizes). The main difference between their paper and ours is that we do not impose

[5] Indeed, Montalvo and Reynal-Querol (2002) show that, within the two-group case, even when group sizes diverge, the two indexes continue to be proportional to each other.

[6] The mixed public-private prize has been used in a different framework by Esteban and Ray (2001). They investigate the group members' ability to overcome the collective action model for different types of prize at stake.

this assumption.[7] In the spirit of ER1999, we investigate the conflict-distribution relationship for varying degrees of publicness of the prize and the implications for the pattern of per capita contributions across groups. In addition, we ask whether the indexes suffer a systematic measurement error relative to the model-based relationship. In this respect, the paper is complementary to Esteban and Ray (2011) as we provide an analytic result explaining some of their numerical simulations.

In the same way as we take seriously the advantages of this class of model, we want to remind its limits. First, we neglect the productive side of the economy. In this sense the relationship between conflict and distribution is a very reduced form. Although the marginal cost of contributing is increasing and captures the rising opportunity cost of devoting resources to a non-productive activity, the prize is exogenous and independent from the level of conflict in the society.[8] Second, we assume a specific ratio contest success function.[9] These modeling choices are driven by reasons of tractability: allowing an arbitrary number of groups in the society complicates the analysis considerably and we had to simplify other aspects of the economy.

In Section 3.1 we describe the model and how it differs from the literature. In Section 3.2 we settle the existence and uniqueness of the equilibrium. In Section 3.3 we analyze the relationship between equilibrium conflict and population distribution.

3.1 Description of the model

Agents. There is a unit mass of individuals distributed over the unit interval, where i indicates the group and k indicates the individual. Individuals are distributed across G groups, each with population n_i, so that $n_i \in (0, 1]$ and $\sum_{i=1}^{G} n_i = 1$.

[7] Their model considers also varying intra-group cohesion and inter-group antagonism. We find that extending our model along those lines would not add additional insight into the model. In fact, they find that intra-group cohesion does not play a role and, for large enough populations, conflict reduces to a weighted average of the fractionalization and polarization indexes (i.e., the Gini coefficient does not matter).

[8] There is a large conflict literature considering the endogeneity of the prize of the contest (Garfinkel and Skaperdas (2007) for an excellent survey).

[9] See Skaperdas (1996) for a general treatment and an axiomatization of contest success functions.

Actions. Society must choose the allocation of a prize. We model this prize directly in terms of the utility individuals receive from it (w_{ik}). We assume that individuals can influence the allocation of the prize by devoting resources into a non-productive activity. The decision process can be interpreted as a lottery, where the probability of receiving the prize is distributed over the population according to a vector of resources. Let $a_{ik} \in \mathbb{R}^+$ denote the resources devoted by individual k in group i. The aggregate amount of resources devoted by the entire population is $A \equiv \sum_{i=1}^{G} \sum_{h \in i} a_{ih}$ (with h indicating the generic individual in group i), where $A \in \mathbb{R}^+$. We will use A as a measure of societal conflict in the non-productive activity.

Timing. The timing is the following: i) all individuals of all groups choose simultaneously their contributions; ii) nature chooses the winning group with probabilities π_i; and iii) the prize is distributed across members of the winning group.

Information. The payoff structure of all individuals is common knowledge.

Payoffs. Let $c(a)$ denote the utility cost of a generic amount of resources. The cost function $c : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is homogeneous across all groups.

Assumption 1. c is continuous, increasing, and twice differentiable with $c(0) = 0$, $c' > 0$, $c'' > 0$ for all $a > 0$, and $\lim_{a \longrightarrow 0^+} c'(a) = c'(0) = 0$.

Define the winning probability of individual k in group i (π_{ik}) as the share of resources devoted by members (indexed by h) of group i :

$$\pi_{ik}(a_{ik}) = \frac{\sum_{h \in i} a_{ih}}{A}, \tag{1}$$

provided $A > 0$. By definition (1), individuals belonging to the same group have the same winning probability: $\pi_{ik} = \pi_{il} = \pi_i \ \forall (k, l) \in i$, $\forall i = 1, .., G$.

Let w_{ik} be the individual benefit from winning the prize. We specify the prize as a mixed private-public good. Let $\lambda \in [0, 1]$ denote the share of publicness of the prize:

$$w_{ik} = w(\lambda, n_i) = \lambda + \frac{1 - \lambda}{n_i}. \tag{2}$$

It is important to specify exactly the nature of the prize. Both the public component (λ) and the private component ($1-\lambda$) are enjoyed exclusively by members of the winning group. The difference between the two is that the per capita benefit associated with the public component is constant, while the one associated with the private component shrinks with group size. The public component can be interpreted in several ways: i) the good is non-excludable (all groups receive it), but only members of the winning group derive utility from it; ii) the good is non-excludable and members of all groups derive utility from it, but members of non-winning groups derive a lower utility than members of the winning group;[10] iii) the good is excludable to members of non-winning groups (and continues to be non-excludable among members of the winning group). With respect to the first two cases, one may think of government policies that are valid for everybody but enjoyed by one particular group.[11] With respect to the last case, one may think of government policies reserved to one particular group.[12] With this caveat in mind, we will hereafter refer to λ as the public component of the prize. A related point is that the prize does not need to be one good with both public and private features. It can also be interpreted as a basket of goods. In this case, λ would be the average share of publicness of the prizes. This interpretation is useful also because the model is the stylized description not necessarily of one contest over one good, but possibly of several contests over several goods, as long as the cleavage that separates the groups remains the same. For simplicity, we assume that the share of publicness of the prize λ is the same across groups. By definition (2), individuals of the same group receive the same benefit in case of capture of the prize: $w_{ik} = w_{il} = w_i \; \forall (k,l) \in i, \; \forall i$.

We assume a utility function for individual k in group i that is linear in the expected benefit from winning the prize net of the cost of contributions:

$$u_{ik}(a_{ik}) = \pi_i(a_{ik}) w_i - c(a_{ik}). \tag{3}$$

[10] In this case w_{ik} constitutes a utility differential.

[11] For example, an eventual extension of public health insurance in the US will benefit those without much more than those with private health insurance. Another example may be the regulation of access to the sea, which applies to any citizen but is enjoyed disproportionally by those living close to the seaside.

[12] For example reservation of political seats for women (Chattopadhyay and Duflo (2004)) or minorities (Pande (2003)).

We assume that individual k in group i chooses his contribution so as to maximize his extended utility function (v_{ik}), which includes the ones of his fellow members:

$$v_{ik}(a_{ik}) = \sum_{l \in i} u_{il}(a_{il}) = u_{ik}(a_{ik}) + \sum_{l \in i, l \neq k} u_{il}(a_{il}). \tag{4}$$

By assuming that individuals maximize this extended utility, we abstract from within-group free-riding. Similar assumptions can be found in Esteban and Ray (1999), Esteban and Ray (2008), Reynal-Querol and Montalvo (2002), and Montalvo and Reynal-Querol (2005).[13] Suppose we were to allow individuals to assign greater weight to their own utility than to that of their fellow members. Then the results would hold as long as they assigned a non-zero weight to their fellow members. Indeed, internalization of fellow members' preferences is thought to be one of the reasons why ethnicity is salient (Alesina and Ferrara (2005)). Even if they did assign zero weight to their fellow members' utilities, all results of the model would resemble the case of pure private goods, which is the main focus of the paper.

To complete the specification of the model, we describe the outcome when $A = 0$. We take this to be an arbitrary vector $\bar{\pi} = (\bar{\pi}_1, .., \bar{\pi}_G)$.[14]

The following table summarizes all variables and functions included in the model.

[13] This assumption can be grounded on one of two theoretical backgrounds: either individual contributions are really determined by a group leader, like in Esteban and Ray (2008), because of coercion or group ideology, or individuals maximize an extended utility, which includes the utility of fellow members (this paper, Esteban and Ray (2011)). ER1999, Reynal-Querol and Montalvo (2002), and Montalvo and Reynal-Querol (2005) assume absence of free-riding, but they leave implicit the theoretical background to support it.

[14] ER1999 provide a similar assumption to complete the specification of their model.

Table 4 - List of the variables in the model.		
a_{ik}	individual contribution of member of individual k in group i	choice variable
n_i	size of group i	exogenous
w_i	utility for any member of group i for outcome $i: \lambda + \frac{1-\lambda}{n_i}$	exogenous
λ	share of publicness of the prize: $\lambda \in [0,1]$	exogenous
π_i	winning probability for any member of group $i: \sum_{i=1}^{G} \pi_i = 1$	endogenous
A	conflict: $A = \sum_{i=1}^{G} \sum_{h \in i} a_{ih}$	endogenous
$c()$	cost of effort $c: \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ and $c(.) : c'(.) > 0, c''(.) > 0$	
a^*	vector of individual contributions $a^* \equiv \left(a^*_{11}, .., a^*_{1n_1}, .., a^*_{G1}, .., a^*_{Gn_G}\right)$	equilibrium
π	vector of winning probabilities $\pi \equiv (\pi_1, .., \pi_G) : \sum_{i=1}^{G} \pi_i = 1$	
N	vector of group sizes $N \equiv (n_1, .., n_G) : \sum_{i=1}^{G} n_i = 1$	

3.2 Agents' behavior and equilibrium

All proofs of the propositions henceforth are presented in the Appendix.

Proposition 1 *Suppose that Assumption 1 holds. Provided $a_{jh} > 0$ for some $j \neq i$, the amount of resources devoted to members of groups i is strictly positive and completely described by the first-order condition (FOC):*

$$\pi_i (1 - \pi_i) w_i (\lambda, n_i) = c' (a_i) a_i. \tag{5}$$

Members of the same group will devote the same per capita contributions $a_{ik} = a_{il} = a_i$ $\forall (k,l) \in i$, $\forall i = 1,..,G$, where a_i denotes the per capita contribution of members of group i. There exists an equilibrium and it is unique.

The first part of Proposition 1 states that the solution to the individual's maximization problem is always interior. Thus, any equilibrium must involve positive contributions by all individuals. Equation (5) provides an intuition of the influence of the mixed prize specification. A larger group implies more fellow members (π_i), but also less opponents ($1-\pi_i$) and, above all, a greater conflict over the private component of the prize, and hence reduced incentives to contribute (smaller benefit w_i). This latter force is more relevant the greater the share of the private component within the prize. This is why we expect both the level and pattern of conflict to vary with the level of this parameter.

The second part of the proposition states that there is one and only one vector of optimal contributions $a^* \equiv (a_1^*, .., a_G^*)$ such that a_{ik}^* solves the maximization of (4) subject to (1), (2), and (3). This implies the existence and uniqueness of equilibrium conflict $A^* = \sum_{i=1}^{G} n_i a_i^*$ and equilibrium winning probabilities $\Pi = (\pi_1^*, .., \pi_G^*)$.

3.3 Conflict and distribution: levels and patterns

In this section we analyze the properties of the model. First, we look at how equilibrium conflict (A) varies with population distribution (N) and the share of publicness of the prize (λ). Since the conflict-distribution relationship for each type of prize $A(\lambda, N)$ is not in an explicit form (see proof of Proposition 1), this is the best way to compare the model to the indexes. Second, we look at how per capita contributions (a_i) vary across groups within a given equilibrium (A fixed).

Recall that our model is an extension of the pure contest version of ER1999 to mixed public-private goods. With respect to our model, their results cover the case of pure public goods ($\lambda = 1$). Throughout the analysis, we refer to their results as a benchmark against which we evaluate ours ($\lambda \in [0, 1)$).

3.3.1 Conflict and distribution: levels

We start our analysis with two general results. First, we investigate the case of two groups. In this case we would expect the uniform distribution to be the global maximum (Tullock 1980).

This is how both the fractionalization (F) and discrete polarization index (Q) behave and what ER1999 find for pure public goods. Second, we investigate the case of an arbitrary number of groups. Over the set of uniform distributions, ER1999 find that equilibrium conflict decreases with the number of groups, provided there are at least two groups to begin with. This is exactly in line with the second property of the discrete polarization index. We investigate whether this continues to be true for all types of goods.

Proposition 2 *Suppose that Assumption 1 holds. Then:*

[1] in the two-group case, equilibrium conflict (A) is maximized at the uniform distribution over the two groups;

[2] over the set of uniform distributions, equilibrium conflict (A) increases with the number of groups up to a threshold $G(\lambda)$, and decreases thereafter. The number of groups maximizing conflict increases as the prize becomes more private $\left(\frac{\partial G(\lambda)}{\partial \lambda} < 0\right)$, and approaches infinity as the prize becomes half public half private $(\lambda = 1/2)$.

Part 1 implies that, in case of $G = 2$, any departure from the uniform distribution, which corresponds to increased population inequality, lowers the level of conflict. The result is consistent both with the fractionalization and discrete polarization indexes and an earlier finding by ER1999.

Part 2 shows that Esteban and Ray's finding is not robust over all types of goods. Most importantly, the conflict-distribution relationship does not resemble the property of the discrete polarization index anymore. On the contrary, for a large set of goods $\left(\lambda \in \left[0, \frac{1}{2}\right]\right)$, conflict increases with the number of groups, thus resembling the second property of the fractionalization index.[15]

Let us now provide some additional results for the special case of pure private goods $(\lambda = 0)$. The next proposition mirrors the list of properties of the fractionalization (F) and polarization (Q) indexes (Section 2.2). First, we identify the distribution that maximizes

[15] Even Esteban and Ray's finding that the symmetric bimodal distribution is the global maximum is not robust to our extension. In fact, we can rule the symmetric bimodal distribution out of the potential candidates for a large set of goods. In order to establish this, it is enough to note that ER's global maximum is a uniform distribution. Since over the set of uniform distributions dissipation is greatest in correspondence of the three-point uniform distribution for $\lambda = \frac{3}{4}$, then the two-points uniform distribution can be ruled out for $\lambda \in \left[0, \frac{3}{4}\right]$.

the level of conflict. Second, we consider the set of uniform distributions. Third, we ask whether there exists a sequence of changes providing unidirectional impacts on conflict, first by looking at the split of a group, then by looking at a generic population transfer from a large to a smaller group. This lets us establish whether the distribution-conflict relationship is monotonic (as opposed to non-monotonic) and local (as opposed to global).

Before the proposition, we spell out two additional assumptions on the cost function that will be useful to identify how generalizable the results are. Let α denote the elasticity of the marginal cost of contribution $c'(a)$ with respect to the contribution itself $a : \alpha(a) = \frac{c''(a)a}{c'(a)}$. We make the following regularity assumptions:

Assumption 2. The cost function is three times differentiable and $c''' \geq -\frac{2c''(a)}{a}$.

Assumption 3. c is three times differentiable and $\alpha'(a) : -[\alpha(a)+1]\alpha(a)+\delta < \alpha'(a)a < [\alpha(a)+1]\alpha(a)-\delta$

The intuition behind both assumptions is that we want the cost function to be "convex enough." They are not very restrictive though. For example, the entire set of iso-elastic cost functions $c(a) = \beta a^{\alpha}$ satisfying Assumption 1 ($\alpha > 1$) satisfies both of them.[16]

We are now ready to present the main finding for pure private goods:

Proposition 3 *Suppose that Assumption 1 holds. Then:*

[1] provided Assumption 2 holds as well, equilibrium conflict (A) *is maximized at the uniform distribution over all groups;*

[2] over the set of uniform distributions, equilibrium conflict (A) *always increases with the number of groups;*

[3] the split of any group increases equilibrium conflict (A)*;*

[4] provided Assumption 3 holds as well, any uniform distribution is always a strict local maximum.

[16] To see this, just note that both the third derivative of an iso-elastic cost-function and the derivative of its elasticity of an iso-elastic function are zero. Assumption 3 is more restrictive than Assumption 2 if $\alpha(a) \in (0,1)$, exactly equal if $\alpha(a) = 1$, and less restrictive if $\alpha(a) > 1$.

Part 1, 2, and 3 coincide with the first three properties of the fractionalization index (F). Part 3 says that the districution-conflict relationship is *local*, in the sense previously defined (Section 2.2). Part 4 says that the conflict-distribution relationship is *monotonic* around the uniform distribution, in the sense previously defined, and it implies that we can not reject the hypothesis that population transfers to smaller groups increase equilibrium conflict (A), or that the distribution-conflict relationship is *monotonic*, in the sense previously defined.[17] In addition, note that the idea that conflict increases as groups become smaller (split) runs against the "divide and conquer" conflict-strategy (ER1999: 397), while it is consistent with the hypothesis that having many independent rent-seeking agencies is worse than having few ones (Shleifer and Vishny (1993)).

On the other hand we know that, in case of pure public goods ($\lambda = 1$), the properties of the distribution-conflict resemble broadly the properties of the discrete polarization index (ER1999). This suggests that the nature of the prize is enough to explain the differences between the fractionalization and discrete polarization indexes, and so that the higher weight assigned to population frequency in the discrete polarization index does not reflect intra-group homogeneity (ER1999) or the sense of identification (Esteban and Ray (1994)), but rather the difference in the prize at stake. Indeed, if we were to include varying intra-group cohesion like Esteban and Ray (2011), we would still find that the properties of the model are close to the Q in case of pure public goods and close to F in case of pure private goods as long as intra-group cohesion was positive.[18]

[17] To see this, consider a sequence of transfers from a uniform distribution over $G-1$ groups to a uniform distribution over G groups. A series of transfers "in the same direction" requires the following steps: first, we split one group so that there is a new group with a very small size; second, we transfer population from all other groups to this small new group. By continuity, the new G-point distribution must have a level of dissipation close to the $G-1$ uniform distribution. From part 3 we know that the level of dissipation corresponding to the new distribution must be greater than the level corresponding to the uniform distribution over $G-1$ groups. From part 4 we know that the uniform distribution over G groups is a local maximum, which means that transfers close to it will be dissipation increasing. We therefore can not reject the hypothesis that each of the transfers affects (increases) the level of dissipation as the one-step change would.

[18] Esteban and Ray (2009) model individuals' extended utility function as a weighted average between one's own utility and the fellow members' utilities. The weight represents the degree of intra-group cohesion. Indeed, if intra-group cohesion were zero, the dissipation-distribution relationship would resemble F for any type of prize

3.3.2 Conflict and distribution: patterns

We will now look at how per capita contributions (a_i) vary across groups within a given equilibrium (A fixed). In particular, we compare per capita contributions (a_i) with the average contribution across the entire population (A). Define the ratio between the two $\left(\frac{a_i}{A}\right)$ as intensity of lobbying. Define activism as any equilibrium such that at least two groups differ in their intensity of lobbying: $a_i \neq a_j$ for some (i, j).

In case of pure public goods, "contests with two groups can never involve activism. On the other hand, contests with more than two groups display activism whenever all groups are not equal-sized, and larger groups always lobby more than smaller groups" (ER1999: 398). This is how results change once we extend the model to mixed public-private goods.

Proposition 4 *Suppose Assumption 1 holds. Then*

[1] in the two-group case, contests involve activism whenever the prize is not a pure public good or the two groups are not equal-sized. In this case, the larger group always lobbies less intensively than the smaller one.

[2] in case of three or more groups and pure private goods, larger groups always lobby less intensively than smaller ones.

Proposition 4.1 illustrates clearly the forces at work described in Section 3.2: a larger group means a greater number of contributions (greater incentive to contribute), but also a smaller opponent (lower incentive to contribute) and lower per capita benefit from the private component of the prize. In case of pure public goods, the latter component does not exist, the first two forces exactly cancel each other out and individuals contribute the same regardless of the population distribution. For all intermediate cases though, the additional incentive created by the private component of the prize plays a role and individuals belonging to the smaller group contribute more than the opponents. In case of an arbitrary number of groups ($G \geq 3$), the second force we listed becomes weaker, yet the third one still dominates. Note that this does not mean that the share of resources devoted by the larger group is smaller than the share of resources devoted by the smaller group. Indeed, the larger group continues to have a greater winning probability (see Lemma 6), but not as much as in the case of pure public goods. Hence, whether we may say that the Pareto-Olson argument plays a role in the model depends on the definition of the latter. According to Esteban and Ray (2001), the Pareto-Olson argument dominates when larger groups have a lower probability of winning than smaller groups, which is not the case here.[19]

Proposition 4 also unveils one difference between the model and the fractionalization index: members of different groups behave differently. This constitutes a new prediction to be tested empirically. It also has some implications for existing empirical evidence:

Proposition 5 *Suppose Assumption 1 holds. Then, in case of pure private goods and a quadratic iso-elastic cost function, the fractionalization index always underestimates the true level of conflict.*

Proposition 5 shows that neglecting the pattern of contributions is not without consequences: the fractionalization index suffers a systematic measurement error.

4 Conclusions

In this paper we asked which population distributions are associated with a high level of conflict and whether the conflict-distribution relationship resembles the fractionalization or the discrete polarization index. In order to answer these questions, we developed a conflict model linking conflict to the distribution of the population across an arbitrary number of groups. The model is an extension of the pure-contest model by Esteban and Ray (1999), who consider only pure public goods and find that the conflict-distribution relationship[20] resembles the discrete polarization index. Here, in contrast, the prize is allowed to vary from pure public goods to pure private goods.

We find that, in case of pure private goods, the conflict-distribution relationship resembles the fractionalization index. This result may explain why cross-country regressions associating

[19] If we relax the assumption of no free-riding, this is not necessarily true (Esteban and Ray 2001).

[20] In their paper they consider the concept of conflict whereas here we consider the concept of dissipation to better interpret the model in light of the empirical stylized facts. However, the modeling strategy is neutral with respect to the concept used.

ethnic diversity with economic performance and likelihood of civil wars are sensitive to the index used to capture the former. To the extent that both reflect competition for the capture of the State, our results suggest that the latter is perceived as a public good in the case of open conflict, while it is perceived as a private good in the cases of lobbying and generalized corruption. It could also be the case that open conflict increases the ability to deliver public goods after the conflict.

The analysis of the per capita contributions across groups suggests that, in case of pure private goods, individuals belonging to smaller groups always contribute more. This suggests that the fractionalization index may systematically underestimate the weight of smaller groups in the creation of conflict. Indeed, we find that, for the special case of quadratic cost functions, the fractionalization index under-estimates the level of conflict. This confirms the pattern in the numerical simulations run by Esteban and Ray (2011) for the case of pure contests, quadratic costs, a large population, and pure private goods. Their simulations are based on random draws for the population vector (over five groups). In this case the divergence between the model-based and index-based levels of conflict appears negligible. Future work should confirm this with real-world data.

References

Ahlerup, P., and O. Olsson (2012): "The roots of ethnic diversity," *Journal of Economic Growth*, 17(2), 71–102.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003): "Fractionalization," *Journal of Economic Growth*, 8(2), 155–194.

Alesina, A., and E. L. Ferrara (2005): "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 43(3), 762–800.

Blattman, C., and E. Miguel (2010): "Civil War," *Journal of Economic Literature*, 48(1), 3–57.

Campos, N. F., and V. S. Kuzeyev (2007): "On the Dynamics of Ethnic Fractionalization," *American Journal of Political Science*, 51(3), 620–639.

Caselli, F., and W. J. I. Coleman (2006): "On the Theory of Ethnic Conflict," .

Chattopadhyay, R., and E. Duflo (2004): "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," *Econometrica*, 72(5), 1409–1443.

Easterly, W., and R. Levine (1997): "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, 112(4), 1203–1250.

Esteban, J., and D. Ray (1994): "On the Measurement of Polarization," *Econometrica*, 62(4), 819–851.

——— (1999): "Conflict and Distribution," *Journal of Economic Theory*, 87(2), 379–415.

——— (2001): "Collective Action and the Group Size Paradox," *American Political Science Review*, 95(03), 663–672.

——— (2008): "On the Salience of Ethnic Conflict," *American Economic Review*, 98(5), 2185–2202.

——— (2011): "Linking Conflict to Inequality and Polarization," *American Economic Review*, 101(4), 1345–74.

Fearon, J. D. (2003): "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth*, 8(2), 195–222.

Garfinkel, M. R., and S. Skaperdas (2007): "Chapter 22 Economics of Conflict: An Overview," in *Handbook of Defense Economics*, ed. by S. Todd, and H. Keith, vol. Volume 2, pp. 649–709. Elsevier.

Michalopoulos, S. (2012): "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102(4), 1508–39.

Montalvo, J., and M. Reynal-Querol (2005): "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review*, 95(3), 796–816.

Pande, R. (2003): "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India," *American Economic Review*, 93(4), 1132–1151.

Reynal-Querol, M., and J. G. Montalvo (2002): "Why Ethnic Fractionalization? Polarization, Ethnic Conflict and Growth," .

Shleifer, A., and R. W. Vishny (1993): "Corruption," *The Quarterly Journal of Economics*, 108(3), 599–617.

Skaperdas, S. (1996): "Contest success functions," *Economic Theory*, 7(2), 283–290.

Spolaore, E., and R. Wacziarg (2009): "The Diffusion of Development," *The Quarterly Journal of Economics*, 124(2), 469–529.

Valsecchi, M. (2010): "Ethnic Diversity, Economic Performance and Civil War," *SWOPEC Working Paper No.433*.

Appendix

Proof. Proposition 1.

Note that maximizing (4) subject to (1), (2) and (3) becomes maximizing

$$\sum_{l\in i}\left[\frac{a_{ik}+\sum\limits_{h\in i,h\neq k}a_{ih}}{A}w_i\right]-c\left(a_{ik}\right)-\sum_{l\in i,l\neq k}c\left(a_{il}\right). \tag{6}$$

Equation (6) is well-defined for every a_{ik} since we have assumed that $a_{jl}>0$ for some $j\neq i$. Differentiation of (6) with respect to a_i provides

$$\begin{aligned}\frac{\partial}{\partial a_{ik}} &= \sum_{l\in i}\left[\frac{A-\left(a_{ik}+\sum\limits_{h\in i,h\neq k}a_{ih}\right)}{A^2}w_i\right]-c'\left(a_{ik}\right)\\ &= \sum_{l\in i}\left[\frac{1}{A}\left(1-\frac{\left(a_{ik}+\sum\limits_{h\in i,h\neq k}a_{ih}\right)}{A}\right)w_i\right]-c'\left(a_{ik}\right),\end{aligned}$$

which, given (1), is

$$=\sum_{l\in i}\left[\frac{1}{A}\left(1-\pi_{ik}\right)w_i\right]-c'\left(a_{ik}\right).$$

Since (by construction) members of the same groups have the same winning probability $(\pi_{ik}=\pi_{il}=\pi_i\forall(k,l)\in i,\forall i)$, then

$$\begin{aligned}\frac{\partial}{\partial a_{ik}} &= \sum_{l\in i}\left[\frac{1}{A}\left(1-\pi_i\right)w_i\right]-c'\left(a_{ik}\right)\\ &= \frac{n_i}{A}\left(1-\pi_i\right)w_i-c'\left(a_{ik}\right).\end{aligned}$$

The end-point restriction on c in Assumption 1 and the observation that the existence of a positive lower bound on the benefit from winning the prize ($w_i\geq 1>0$) ensure that the solution to the maximization problem is interior (the optimality condition, FOC, must hold with equality):

$$a_{ik}^{*}:\frac{n_i}{A}\left(1-\pi_i\right)w_i=c'\left(a_{ik}^{*}\right). \tag{7}$$

Since the expected benefit $(\pi_i w_i)$ is strictly concave in a_{ik} and Assumption 1 ensures that the cost function is strictly convex, the individual utility function is strictly concave, which means that equation (7) is also sufficient to define the solution.

Since this is true for every member of group i, then $a_{ik}^{*}=a_{il}^{*}=a_i^{*}\ \forall\left(k,l\right)\in i,\ \forall i$. This lets us rewrite equation (7) as equation (5). It also lets us rewrite (1) as $\pi_i=\frac{n_i a_i^{*}}{A}$. In order to establish the existence and uniqueness of the equilibrium, define a function $\phi:[0,1]^2\times\mathbb{R}^{+}\rightarrow\mathbb{R}$ such that the single element $\phi\left(\pi_i,A,n_i\right)$ is defined by the first-order derivative of the maximization problem in terms of winning probability, conflict, and group size $\left(a_i=\frac{\pi_i A}{n_i}\right)$:

$$\begin{aligned}&\frac{n_i}{A}\left(1-\pi_i\right)w_i-c'\left(\frac{\pi_i A}{n_i}\right)\\=\;&\phi\left(\pi_i,A,n_i\right).\end{aligned}$$

Redefine the equilibrium as any combination of winning probabilities $\pi^{*}=(\pi_1^{*},..,\pi_G^{*})$ and total effort A^{*}, such that $\phi\left(\pi_i^{*},A^{*},n_i\right)=0\ \forall i$, and $\sum_{i=1}^{G}\pi_i^{*}=1$.

The determination of the equilibrium can be shown in two steps: first, by making reference to the individual FOC; second, by making reference to the probability consistency condition $\left(\sum_{i=1}^{G}\pi_i=1\right)$.

Suppose A (and N) is fixed, and consider the behavior of the first derivative $\phi\left(\pi_i,A,n_i\right)$ as the winning probability (π_i) varies along its domain $[0,1]$:

- $\frac{\partial\phi(\pi_i,A,N_i)}{\partial\pi_i}=-\frac{n_i}{A}w_i-\frac{A}{n_i}c''\left(\frac{\pi_i A}{n_i}\right)<0$ (strictly decreasing);
- $\lim_{\pi\longrightarrow0^{+}}\phi\left(\pi_i,A,n_i\right)=\frac{n_i}{A}w_i>0$;
- $\lim_{\pi\longrightarrow1^{-}}\phi\left(\pi_i,A,n_i\right)=-c'\left(\frac{A}{n_i}\right)<0$.

The intermediate value theorem ensures the existence and uniqueness of a winning probability satisfying the equilibrium condition: $\exists!\pi_i^{*}:\phi\left(\pi_i^{*},A,n_i\right)=0$. This value can be thought of as a function depending on the remaining variables: $\pi_i^{*}=\pi\left(A,n_i\right)$.

Aggregate consistency requires the sum of these winning probabilities to equal unity: $\sum_{i=1}^{G} \pi(A, n_i) = 1$. Suppose that N is fixed and consider the behavior of the sum of winning probabilities $\left(\sum_{i=1}^{G} \pi(A, n_i)\right)$ as total conflict (A) varies along its domain $[0, +\infty)$. Since we have not derived an explicit expression for the equilibrium winning probability, we refer to the implicit function theorem to study it. Rewrite the FOC function $\phi_i \equiv \phi(\pi_i, A, n_i)$. Then we know that

$$\frac{\partial \phi_i}{\partial \pi_i} \frac{d\pi(A, n_i)}{dA} + \frac{\partial \phi_i}{\partial A} = 0,$$

which means that

$$\frac{d\pi(A, n_i)}{dA} = -\frac{\frac{\partial \phi_i}{\partial A}}{\frac{\partial \phi_i}{\partial \pi_i}}.$$

Since $\frac{\partial \phi_i}{\partial A} = -\frac{n_i}{A^2}(1-\pi_i) w_i - \frac{\pi_i}{n_i} c''\left(\frac{\pi_i A}{n_i}\right) < 0$ and $\frac{\partial \phi_i}{\partial \pi_i} = -\frac{n_i}{A} u - \frac{A}{N_i} c''\left(\frac{\pi_i A}{n_i}\right) < 0$, then

$$\frac{d\pi(A, n_i)}{dA} < 0 \forall i,$$

which implies

$$\sum_{i=1}^{G} \frac{d\pi(A, n_i)}{dA} < 0 \Longrightarrow \frac{d\left[\sum_{i=1}^{G} \pi(A, n_i)\right]}{dA} < 0.$$

Again, we derive the behavior of this function as total conflict approaches the limits of its domain. In order to do so, we focus on the single winning probability $\pi(A, n_i)$. In order to determine the behavior of the winning probability for any member of group i as total conflict shrinks to zero, fix the winning probability and consider the behavior of the first derivative ϕ as total conflict shrinks to zero:

$$\begin{aligned}
\lim_{A \longrightarrow 0^+} \phi(\pi_i, A, n_i) &= \lim_{A \longrightarrow 0^+} \frac{n_i}{A}(1-\pi_i) w_i - \lim_{A \longrightarrow 0^+} c'\left(\frac{\pi_i A}{n_i}\right) \\
&= \infty - c'(0) = \infty.
\end{aligned}$$

For the first-order condition $(\phi_i = 0)$ to continue to hold, the winning probability must approach unity as total conflict $\left(\lim_{A \longrightarrow 0^+} \pi(A, n_i) = 1\right)$. This implies that the sum of winning probabilities will exceed unity: $\lim_{A \longrightarrow 0^+} \left[\sum_{i=1}^{G} \pi(A, n_i) \right] = G(> 1)$. In order to determine the behavior of the winning probability for any member of group i as total conflict increases to infinity, fix the winning probability and consider the behavior of the first derivative ϕ as total conflict increases to infinity:

$$\begin{aligned} \lim_{A \longrightarrow +\infty} \phi(\pi_i, A, n_i) &= \lim_{A \longrightarrow +\infty} \frac{n_i}{A}(1 - \pi_i) w_i - \lim_{A \longrightarrow +\infty} c' \left(\frac{\pi_i A}{n_i} \right) \\ &= 0 - \infty = -\infty. \end{aligned}$$

For the first-order condition ($\phi_i = 0$) to continue to hold, the winning probability must shrink to zero $\left(\lim_{A \longrightarrow +\infty} \pi(A, n_i) = 0 \right)$. This implies that the sum of winning probabilities will shrink to zero as well: $\lim_{A \longrightarrow +\infty} \left[\sum_{i=1}^{G} \pi(A, n_i) \right] = 0$.

Given the last three results $\left(\frac{d\left[\sum_{i=1}^{G} \pi(A, n_i) \right]}{dA} < 0, \ \lim_{A \longrightarrow 0^+} \left[\sum_{i=1}^{G} \pi(A, n_i) \right] = G, \ \lim_{A \longrightarrow +\infty} \left[\sum_{i=1}^{G} \pi(A, n_i) \right] = 0 \right)$, the intermediate value theorem ensures the existence and uniqueness of a value of total conflict satisfying the equilibrium condition: $\exists! A^* : \sum_{i=1}^{G} \pi(A^*, n_i) = 1$. Such value can be thought of as depending on the vector of group sizes $N = (n_1, .., n_G)$: $A^* = A(N)$.

In summary, for any vector of group sizes N, there is one and only one level of total effort and vector of winning probabilities satisfying the equilibrium conditions. ■

Proof. Proposition 2.

Part 1 means that the uniform distribution is the strict global maximum. Since there are only two groups (1,2), and their sizes (n_1, n_2) must add to unity, we can just re-define their sizes as $n_1 = n$ and $n_2 = 1 - n$. The conflict function $A(N)$ can be redefined accordingly, $A(n)$. Redefine the group's winning probability, $\Pi_i(n) \equiv \pi_i(A(n), n)$. Since the probabilities of winning must also add to unity, then $\Pi_1 = \Pi$ and $\Pi_2 = 1 - \Pi$. Redefine the first-order derivative accordingly: $\phi(\pi, A, n) = \Phi(\Pi(n), A(n), n) \equiv \Phi_1$. The first-order derivative of this function with respect to n is

$$\frac{d\Phi(\Pi, A, n)}{dn} = \frac{\partial \Phi_1}{\partial \Pi} \frac{d\Pi}{dn} + \frac{\partial \Phi_1}{\partial A} \frac{dA}{dn} + \frac{\partial \Phi_1}{\partial n} = 0. \tag{8}$$

Explicit the derivative of the winning probability with respect to $n:$

$$\frac{d\Pi}{dn} = -\frac{\frac{\partial \Phi_1}{\partial A}\frac{dA}{dn} + \frac{\partial \Phi_1}{\partial n}}{\frac{\partial \Phi_1}{\partial \Pi}}.$$

Since population is normalized to unity, an infinitesimal change in the size of group 1 (n) directly affects also the size of group 2 $(1-n)$. Let the first-order derivative for the generic member of group 2 be $\Phi_2 \equiv \Phi(1-\Pi, A, 1-n)$. There will be another direct and indirect effect to consider. However, we know that the sum of winning probabilities must be equal unity before and after the shift. Therefore, the two aggregate changes in winning probabilities must compensate each other: $\sum_{i=1}^{2} \frac{d\Pi_i}{dn} = 0$. Then we can explicit the total derivative of conflict A with respect to the population parameter $\left(\frac{dA}{dn}\right)$:

$$\frac{dA}{dn} = -\frac{\sum_{i=1}^{2}\left[\frac{\partial \Phi_i/\partial n}{\partial \Phi_i/\partial \Pi}\right]}{\sum_{i=1}^{2}\left[\frac{\partial \Phi_i/\partial A}{\partial \Phi_i/\partial \Pi}\right]}.$$

Let α denote the elasticity of the marginal cost of contribution $c'(a)$ with respect to the contribution itself $a : \alpha(a) = \frac{c''(a)a}{c'(a)}$. Let θ denote the ratio between the share of publicness of the prize (λ) and the benefit from winning the prize (w):

$$\theta = \frac{\lambda}{w} = \frac{\lambda}{\lambda + (1-\lambda)/n}. \tag{9}$$

The two initial first-order derivatives Φ_i are $\Phi_1 = \frac{n}{A}(1-\Pi)w(n) - c'\left(\frac{\Pi A}{n}\right)$ and $\Phi_2 = \frac{1-n}{A}\Pi w(1-n) - c'\left(\frac{(1-\Pi)A}{(1-n)}\right)$. Differentiation of these two expressions and some manipulation provides the following expression, where $\alpha_1 = \alpha\left(\frac{\Pi A}{n}\right)$ and $\alpha_2 = \alpha\left(\frac{(1-\Pi)A}{(1-n)}\right)$:

$$\frac{dA}{dn} = \frac{A}{n(1-n)} \frac{(\alpha_1+\theta_1)(1-n)\left[\Pi\alpha_2 + (1-\Pi)\right] + (\alpha_2+\theta_2)n\left[(1-\Pi)\alpha_1 + \Pi\right]}{(1-2n)(\alpha_1\alpha_2 - 1)}.$$

It follows that

$$sign\left\{\frac{dA}{dn}\right\} = sign\left\{(1-2n)\right\},$$

which means that $A(n)$ is increasing in n for $n \in \left(0, \frac{1}{2}\right]$ and decreasing afterwards. Therefore, $A(n)$ attains its maximum at $n = \frac{1}{2}$, which corresponds to the uniform distribution over the two groups. Thus, part 1 is established.

To establish part 2, note that, as we restrict our attention to uniform distributions $(n_i = n\ \forall i)$, the maximization problem becomes identical for individuals across all groups. Per capita contributions are identical $(a_i = a_j = a\ \forall i, j)$ and so are winning probabilities $(\pi_i = \pi = n\ \forall i)$. Given the normalization of total population to unity, equilibrium contributions will also equal total conflict $(a = A)$. Equation (5) reduces to

$$n(1-n)\,w(n) = c'(A)\,A.$$

Define a new function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that $f(a) \equiv c'(a)a$. This let us rewrite the previous equality as

$$n(1-n)\,w(n) = f(A).$$

Assumption 1 ensures that $f(.)$ is strictly increasing: $f'(A) = c''(A)\,A + c'(A) > 0$. This means that f is invertible and the conflict-maximizing problem reduces to maximizing the LHS:

$$\begin{aligned} & \max_n \{n(1-n)\,w(n)\} \\ = & \max_n \left\{ n(1-n)\left(\lambda + \frac{1-\lambda}{n}\right)\right\} \\ = & \max_n \{n(1-n)\,\lambda + (1-n)(1-\lambda)\}, \end{aligned}$$

$$FOC : (1-2n)\,\lambda - (1-\lambda) \leq 0\ (= 0\ if\ n > 0).$$

If the share of publicness of the prize (λ) is equal to or smaller than $\frac{1}{2}$, the solution is corner $(n = 0)$. Otherwise the solution is interior and equal to

$$n = 1 - \frac{1}{2\lambda} \equiv n(\lambda).$$

The number of groups corresponding to these solutions is $G(\lambda) = \frac{1}{n(\lambda)}$, which means that $G(\lambda) = +\infty \ \forall\lambda \in \left[0, \frac{1}{2}\right]$ and $G(\lambda) = \frac{2\lambda}{2\lambda - 1}$. In particular, note that

$$\frac{\partial G(\lambda)}{\partial \lambda} < 0 \quad \forall\lambda \in \left(\frac{1}{2}, 1\right]$$

and $G(1) = 2$. Thus, part 2 is established. ■

Proof. Proposition 3.

Reconsider equation (5) in case of pure private goods ($\lambda = 0$):

$$\pi_i (1 - \pi_i) \frac{1}{n_i} = c'(a_i) a_i. \tag{10}$$

Define a new function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that $f(a) \equiv c'(a)a$. This lets us rewrite equation (8) as

$$\pi_i (1 - \pi_i) = n_i f(a_i).$$

Aggregate over groups to obtain

$$\sum_{i=1}^{G} [\pi_i (1 - \pi_i)] = \sum_{i=1}^{G} [n_i f(a_i)].$$

Assumption 1 ensures that $f(.)$ is strictly increasing: $f'(a) = c''(a) a + c'(a) > 0$. Assumption 2 ensures that $f(.)$ is convex: $f''(a) = c'''(a)a + 2c''(a) \geq 0$. This lets us use the Jensen inequality theorem: $\sum_{i=1}^{G} [n_i f(a_i)] \geq f\left(\sum_{i=1}^{G} n_i a_i\right)$, where $f\left(\sum_{i=1}^{G} n_i a_i\right) = f(A)$. In turn, we know that

$$\sum_{i=1}^{G} [\pi_i (1 - \pi_i)] \geq f(A).$$

Maximizing the LHS subject to the constraint that the sum of winning probabilities must be equal to unity $\left(\sum_{i=1}^{G} \pi_i = 1\right)$ provides the uniform distribution $\bar{\pi} = (\pi, .., \pi)$:

$$G[\pi (1 - \pi)] \geq \sum_{i=1}^{G} [\pi_i (1 - \pi_i)],$$

with equality only if $\pi_i = \pi \ \forall i$.

From the proof of existence and uniqueness of the equilibrium, we know that there is only one population vector that corresponds to the uniform winning probability vector, and that it is the uniform population vector $\bar{N} = (n, .., n)$. Let A^* denote the conflict level corresponding to this maximum. Then we know that

$$f(A^*) \geq f(A).$$

Since f is strictly increasing, this implies $A^* \geq A$, with equality if and only if $N = \bar{N}$. Thus, part 1 is established.

Part 2 is a special case of Proposition 2 (part 2) and hence it is already established.

To establish part 3, rephrase it without loss of generality as "any merger lowers equilibrium conflict." The following definition will be used frequently throughout the proof. In order to clarify the exposition, we drop the subscripts. Define the subjective share of publicness of the prize (θ) as the ratio between the share of publicness of the prize (λ) and the benefit from winning the prize (w):

$$\theta = \frac{\lambda}{w} = \frac{\lambda}{\lambda + (1-\lambda)/n}. \tag{11}$$

The following lemma describes properties that will be needed in the proofs of Propositions 5, 6 and 7. ■

Lemma 6 *Suppose that Assumption 1 holds. Then*

[1] the function $\pi(.)$ is strictly increasing and twice continuously differentiable;

[2] provided $\lambda = 0$, $\left(\frac{\pi}{n}\right)$ is strictly decreasing;

[3] provided $\lambda = 0$, if $(a, b) >> 0$, then $\pi(a+b) < \pi(a) + \pi(b)$.

Proof. Recall that $\pi(.)$ is implicitly defined by equation (5), which we can rewrite in terms of (π, A, n) as

$$\frac{n}{A}(1-\pi)\,w(n) = c'\left(\frac{\pi A}{n}\right). \tag{12}$$

Set A fixed and differentiate equation (10) with respect to n to obtain

$$\pi'(n) = \frac{\pi}{n}\frac{(1-\pi)[\alpha(a)+\theta]}{(1-\pi)\alpha(a)+\pi}. \tag{13}$$

Assumption 1 ensures $\alpha(a) > 0 \ \forall a > 0$. Therefore, $\pi'(.) > 0 \ \forall n > 0$. Thus, part 1 is established.

Using (11) we can derive the derivative of the ratio between winning probability and group size $\left(\frac{\pi}{n}\right)$ with respect to size (n):

$$\frac{\partial\left(\frac{\pi}{n}\right)}{\partial n} = \frac{\pi'(n)}{n}\frac{\theta-(\theta+1)\pi}{(1-\pi)[\alpha(a)+\theta]}. \tag{14}$$

Equation (12) shows that

$$sign\left\{\frac{\partial\left(\frac{\pi}{n}\right)}{\partial n}\right\} = sign\left\{\theta-(\theta+1)\pi\right\}.$$

In case of pure private goods $\theta = 0$, so $\frac{\partial\left(\frac{\pi}{n}\right)}{\partial n} < 0 \ \forall n, \forall \lambda$. Thus, part 2 is established.

Consider $(a,b) >> 0$. From part 2 we know that $\frac{\pi(a+b)}{a+b} < \frac{\pi(a)}{a}$ and $\frac{\pi(a+b)}{a+b} < \frac{\pi(b)}{b}$. It follows that

$$\begin{aligned}\pi(a+b) &= \frac{a+b}{a+b}\pi(a+b) = a\frac{\pi(a+b)}{a+b} + b\frac{\pi(a+b)}{a+b} \\ &< a\frac{\pi(a)}{a} + b\frac{\pi(a)}{a} = \pi(a)+\pi(a).\end{aligned}$$

Thus, part 3 is established. ■

Proof. We return to the main proof.

Sort groups according their winning probabilities (π_i). Consider any sub-set M of the G groups. From Lemma 6 we know that

$$\pi\left(A, \sum_{i\in M} n_i\right) < \sum_{i\in M}\pi(A, n_i).$$

Add the winning probabilities of all remaining groups $(j \neq M)$, evaluated at the initial level of conflict A :

$$\pi\left(A,\sum_{i\in M}n_i\right)+\sum_{j\neq M}\pi\left(A,n_j\right)<\sum_{i\in M}\pi\left(A,n_i\right)+\sum_{j\neq M}\pi\left(A,n_j\right)=1.$$

For the sum of winning probabilities to equal unity also in the final distribution, the level of conflict must decrease: $A'<A$. Therefore any merger must decrease the level of conflict (and any split must increase it). Thus, part 3 is established.

To establish part 4, consider the G-point uniform distribution $\bar{N}_G=(n,n,..)$. Call the corresponding level of conflict $\bar{A}_G$. Set $\bar{A}_G$ fixed and differentiate (11) with respect to n. After some manipulation, we obtain

$$\begin{aligned}\pi''(n) &= \frac{\left[\pi'(n)\right]^2}{\pi(1-\pi)\left[\alpha(a)+\theta\right]\left[(1-\pi)\alpha(a)+\pi\right]} \\ &\quad \left\{\begin{array}{c}\left[\theta-(\theta+1)\pi\right]\left[(1-\pi)\alpha(a)+\pi\right]+\frac{\theta(1-\theta)}{\alpha(a)+\theta}\left[(1-\pi)\alpha(a)+\pi\right]^2+ \\ -\left[\alpha(a)+\theta\right]\pi-\frac{\alpha'(a)a}{\alpha(a)+\theta}\left[\theta-(\theta+1)\pi\right]^2\end{array}\right\},\end{aligned}$$

where $\theta=\theta(\lambda,n)$.

Define the expression in curly brackets as $\varphi(\lambda,n)$. Clearly, $sign\{\pi''(n)\}=sign\{\varphi(\lambda,n)\}$.

The case of purely private goods corresponds to setting $\lambda=0$, which means $\theta(0,n)=0$. By substitution, we find

$$\begin{aligned}\varphi(0,n) &= \left\{-\pi\left[(1-\pi)\alpha(a)+\pi\right]-\alpha(a)\pi-\frac{\alpha'(a)a}{\alpha(a)}\pi^2\right\} \\ &= -\pi\left\{(1-\pi)\alpha(a)+\pi+\alpha(a)+\frac{\alpha'(a)a}{\alpha(a)}\pi\right\} \\ &= -\pi\left\{2(1-\pi)\alpha(a)+\left[\alpha(a)+1+\frac{\alpha'(a)a}{\alpha(a)}\right]\pi\right\}.\end{aligned}$$

By Assumption 3, the equation becomes

$$\varphi(0,n)=-\pi\left\{2(1-\pi)\alpha(a)+\delta\pi\right\},$$

where $\frac{\alpha'(a)a}{\alpha(a)}=-\left[\alpha(a)+1\right]+\delta$. Since $\delta>0$, then $\varphi(0,n)<0\ \forall n>0$. The winning probability $\pi(.)$ is locally strictly concave in an open neighborhood around the point combination $\left(\bar{A}_G(\lambda), n\right)$. Pick any G-point non-uniform distribution $\tilde{N}_G = (\tilde{n}_{G1}, .., \tilde{n}_{GG})$ such that the combination $\left(\bar{A}_G(\lambda), \tilde{n}_{Gi}\right)$ lies in the open neighborhood of $\left(\bar{A}_G(\lambda), n\right)$ for every i. By local strict concavity and the equilibrium condition $\sum_{i=1}^{G} \pi(A, n_i) = 1$,

$$1 = G\pi\left(\bar{A}_G(\lambda), n\right) > \sum_{i=1}^{G} \pi\left(\bar{A}_G(\lambda), \tilde{n}_{Gi}\right).$$

Let $\tilde{A}_G(\lambda)$ be the equilibrium conflict associated with $\tilde{N}_G$. Recall that $\pi(.)$ is strictly decreasing in A : $\frac{d\pi(A_G(\lambda), \tilde{n}_{Gi})}{dA} < 0$ $\left(\text{as well as } \frac{d\left[\sum_{i=1}^{G} \pi(A_G(\lambda), \tilde{n}_{Gi})\right]}{dA} < 0\right)$ $\forall i$. This, joint to the previous inequality, implies $\bar{A}_G(\lambda) > \tilde{A}_G(\lambda)$. Thus, part 4 is established. ■

Proof. Proposition 4.

Note that the ratio between a group's per capita contribution and average contribution $\left(\frac{a_i}{A}\right)$ is exactly equal to the ratio between probability of winning and group size $\left(\frac{\pi_i}{n_i}\right)$.

Consider the case $G = 2$. Let n be the size of group 1 and $(1-n)$ the size of group 2. Let π be the winning probability of group 1 and $(1-\pi)$ the winning probability of group 2. Consider the ratio between the FOC of two individuals belonging to different groups:

$$\begin{aligned} \frac{c'(a_1)\, a_1}{c'(a_2)\, a_2} &= \frac{w(\lambda, n)}{w(\lambda, 1-n)} \\ &= \frac{1-n}{n} \frac{\lambda n + 1 - \lambda}{\lambda(1-n) + 1 - \lambda}. \end{aligned}$$

If the RHS is greater then unity, group 1 lobbies more intensively than group 2. If the two groups have equal size $\left(n = \frac{1}{2}\right)$, the RHS is equal to unity, which means absence of activism. Consider the general case

$$\frac{1-n}{n} \frac{\lambda n + 1 - \lambda}{\lambda(1-n) + 1 - \lambda} \geq 1.$$

After some manipulations, we find that

$$(1-2n)(1-\lambda) \geq 0.$$

If the good is purely public $(\lambda = 1)$, then the inequality is satisfied for any value of n. If the good is intermediate or purely private $(\lambda < 1)$, then the inequality is satisfied for $n \leq \frac{1}{2}$ (with strict inequality if $n < \frac{1}{2}$). This means that the bigger group lobbies less intensively than the smaller one.

Consider the case $G \geq 3$ and the special case $\lambda = 0$. Sort groups with respect to their size. Recall from Lemma 6 that the ratio $\left(\frac{\pi}{n}\right)$ is decreasing in n, which means that bigger groups lobby less intensively than smaller ones. ■

Proof. Proposition 5.

In case of pure private goods $(\lambda = 0)$, and an iso-elastic cost function $c(a) = \frac{a^2}{2}$, we get $A^2 = \sum_{i=1}^{G} [n_i (1 - \pi_i)]$. Recall the formula for the fractionalization index: $F = \sum_{i=1}^{G} [n_i (1 - n_i)]$. Proposition 8 says that the conflict A^2 is always greater than fractionalization $(F) : A^2 > F$. This can be written as

$$\begin{aligned} A^2 - F &= \sum_{i=1}^{G} [n_i (1 - \pi_i)] - \sum_{i=1}^{G} [n_i (1 - n_i)] \qquad (15) \\ &= \sum_{i=1}^{G} \{[(1 - \pi_i) - (1 - n_i)] n_i\} \\ &= \sum_{i=1}^{G} [(n_i - \pi_i) n_i]. \end{aligned}$$

Sort groups so that $n_1 \leq .. \leq n_G$. Since $\pi'(n) > 0$, the same sorting applies to winning probabilities: $\pi_1 \leq .. \leq \pi_G$. Lemma 6 ensures that the ratio $\left(\frac{\pi}{n}\right)$ is decreasing in n: $\frac{\pi_1}{n_1} \geq .. \geq \frac{\pi_G}{n_G}$. Since $\frac{\pi_i}{n_i} = \frac{a_i}{A}$ and A is a weighted average of per capita contributions (given that population is normalized to unity), then $A \in [a_1, a_G]$ (with equality only in case of uniform distribution). This implies that $\exists! n^* \in [n_1, n_G]$[21] $: \frac{\pi(n^*)}{n^*} = 1$, or, $\pi(n^*) = n^*$. n^* divides the groups in the following way: $\pi_i > n_i \ \forall i \in \{n_i < n^*\}$; $\pi_i < n_i \ \forall i \in \{n_i > n^*\}$; $\pi_i = n_i$

[21] There is only one case where $n^* = n_1$ or $n^* = n^G$. It corresponds to the uniform distribution $(n_1 = .. = n_G)$.

$\forall i \in \{n_i = n^*\}$. Hence,

$$\begin{aligned} \sum_{i\in\{n_i<n^*\}} [(n_i - \pi_i)\, n_i] &< 0; \\ \sum_{i\in\{n_i>n^*\}} [(n_i - \pi_i)\, n_i] &> 0; \\ \sum_{i\in\{n_i=n^*\}} [(n_i - \pi_i)\, n_i] &= 0. \end{aligned}$$

Define $\hat{n} : \hat{n} \in \{n_i < n^*\} \wedge n_i < \hat{n}\ \forall i \in \{n_i < n^*\}$. This lets us establish a lower bound to the first subset:

$$\sum_{i\in\{n_i<n^*\}} [(n_i - \pi_i)\, \hat{n}] \le \sum_{i\in\{n_i<n^*\}} [(n_i - \pi_i)\, n_i]\,.$$

Define $\check{n} : \check{n} \in \{n_i > n^*\} \wedge n_i > \check{n}\ \forall i \in \{n_i > n^*\}$. This lets us establish a lower bound to the first subset:

$$\sum_{i\in\{n_i>n^*\}} [(n_i - \pi_i)\, \check{n}] \le \sum_{i\in\{n_i>n^*\}} [(n_i - \pi_i)\, n_i]\,.$$

In addition, note that the two group size thresholds are ordered: $\hat{n} < \check{n}$.

Disaggregate equation (13) with respect to the subgroups and use these inequalities:

$$\begin{aligned} & \sum_{i\in\{n_i<n^*\}} [(n_i - \pi_i)\, n_i] + \sum_{i\in\{n_i>n^*\}} [(n_i - \pi_i)\, n_i] \\ \ge\ & \sum_{i\in\{n_i<n^*\}} [(n_i - \pi_i)\, n_i] + \check{n} \sum_{i\in\{n_i>n^*\}} [(n_i - \pi_i)] \\ \ge\ & \hat{n} \sum_{i\in\{n_i<n^*\}} (n_i - \pi_i) + \check{n} \sum_{i\in\{n_i>n^*\}} (n_i - \pi_i) \\ >\ & \hat{n} \sum_{i\in\{n_i<n^*\}} (n_i - \pi_i) + \hat{n} \sum_{i\in\{n_i>n^*\}} (n_i - \pi_i) \\ =\ & \hat{n} \sum_{i=1}^{G} (n_i - \pi_i) = 0 \end{aligned}$$

The last equality comes from the fact that $\sum_i \pi_i = \sum_i n_i \Rightarrow \sum_i (\pi_i - n_i) = 0$. Thus, we have established that $A^2 - F > 0$, which proves Proposition 5. ■

Paper 2

Land Property Rights and International Migration:

Evidence from Mexico

Michele Valsecchi*

University of Gothenburg

First version: April 2010

Updated: 8 September 2012

Abstract

In this paper we ask whether there is a relationship between land property rights and international migration. In order to identify the impact of property rights, we consider a country-wide land certification program that took place in Mexico in the 1990s. Our identification strategy exploits the timing of the program and the heterogeneity in farmers' eligibility for the program. Comparing eligible and ineligible households, we find that the program increased the likelihood of having one or more members abroad by 12 percent. In terms of number of migrants, our coefficient estimates explain 31 percent of the 1994-1997 increase in migrants from ejido areas and 16-18 percent of the increase from the entire Mexico. We contribute to the current debate on the determinants of Mexican emigration (Hanson 2006, Hanson and McIntosh 2009, Hanson and McIntosh 2010). Consistent with our theoretical model, the impact is strongest for households without a land will.

*I am grateful to Jean-Marie Baland, Ryan Bubb, Alain de Janvry, Gordon Hanson, Karen Macours, Masayuki Kudamatsu, Mariapia Mendola, Ola Olsson, Elisabeth Sadoulet, Måns Soderbom, Alessandro Tarozzi, Yanos Zylberberg, and seminar participants at the University of Gothenburg, NEUDC 2011 at Yale, LACEA 2010 Annual Meeting in Medellin, 3rd International Conference on Migration and Development at Paris School of Economics, EUDN 2010 Workshop in Clermont-Ferrand, and NCDE 2010 Conference in Helsinki. This paper came about following a project with Mariapia Mendola. I wish to thank Elisabeth Sadoulet and Alain de Janvry for providing access to the data. Research funding was provided by Adlerbertska Studentbostadsstiftelsen. Previous versions of this paper were circulated with the title "Land Certification and International Migration: Evidence from Mexico". Email address: michele.valsecchi@economics.gu.se. All errors are my own.

Key words: Property rights; land titling; land reform; land inheritance; international migration.

JEL Classification codes: F22, D23, Q15.

1 Introduction

From 1990 to 2005, the share of Mexicans in the United States increased from 5.2 percent to 10.2 percent (Hanson (2010)). During the same period, remittances from the US to Mexico rose from US$2.5 billion to US$21.7 billion, with an average of US$7.5 billion, or 59% of the net FDI (World Bank (2010)). Mexico is the main source of both legal and illegal immigration to the US. In 2004, 56 percent of the 10.3 million Mexicans in the US were there illegally (Passel (2005)). Hence, illegal immigration causes a huge pressure on the US government to limit border crossing (Hanson and Spilimbergo (1999)), drives the political fortunes of US Governors (Hanson (2005)) and stands high on the agenda of every US presidential candidate. Understanding what drives this migration flow is critical for any assessment of future patterns and policy design (Hanson (2006)).

Although recent contributions attribute a large share of this rise in migration to demographic factors (Hanson and McIntosh (2009), Hanson and McIntosh (2010)), much remains to be understood. In the 1990s, the Mexican government implemented various policies that may have affected migration, yet we lack rigorous econometric evidence in this respect (Hanson (2006)). We contribute to the literature by showing that changes in land property rights in the 1990s did affect migration to the US.

The research questions are, is there a relationship between land property rights and Mexico-US migration? If there is, do better defined property rights slow down or speed up migration flows?

In order to identify the impact of property rights on migration behavior, we make use of the land certification program *Procede*, which was implemented throughout the 1990s and targeted all ejido land in the country. Ejidos are areas of land allocated in usufruct to groups of farmers, called ejidatarios, and cover about 60 percent of all agricultural land in the country (Velez (1995)). *Procede* provided households with certificates for their housing plot, their individuals plots, and their right to use the common land. By providing certainty over their rights, the certificates may have led households to relocate their labor supply in favor of off-farm activities, like migration. In order to account for potential omitted variable bias, we exploit program timing and households' eligibility for the program. Comparing eligible and ineligible households, we find that the program increased the likelihood of having one or more members abroad by 12 percent. In terms of number of migrants, our coefficient estimates explain 31 percent of the 1994-1997 increase in Mexican migrants from ejido areas and 16-18 percent of the increase from the entire Mexico.

The paper also contributes to the literature on land property rights and titling programs, and to the literature on international migration. Concerning the latter, in his recent survey, Hanson (2010) argues that, notwithstanding the recent rise in global migration, it is very challenging to reconcile the level of global migrants (about 3 percent of the global population) with large and persistent wage differentials across countries. This is even more puzzling in the case of Mexico, where borders are porous and illegal migration is widespread. Hanson (2006) calculates that at the existing wage rates (confirmed by Rosenzweig (2007)), it takes less than two months for a migrant with 5-8 years of education to recoup the costs of crossing the border.

There are two sets of explanations. First, cross-country wage differentials may be lower than the average earning differences if migrants' self-selection is positive. This may not apply to Mexico as Chiquiar and Hanson (2005) find that selection there is intermediate.[1] Second, there must be large unobserved costs of migrating other than the cost of crossing the border. However, rather than identifying these costs, the literature has focused on the cost-mitigating role of networks at the destination (see Munshi (2003) and McKenzie and Rapoport (2010) and references therein). The present paper contributes to this literature by identifying a strong yet neglected determinant of migration: tenure (in)security. Tenure insecurity may have induced household members to stay home in order not to lose their land inheritance. Moreover, it may have reduced the incentive to use migration as a self-funding strategy to send money back home (Woodruff and Zenteno (2007), Yang (2008), Mendola (2008)).

We also contribute to the literature on land titling programs. In the last decade, research has mainly aimed at estimating the impact on investments (see Pande and Udry (2006),

[1] Evidence is not conclusive though; see Orrenius and Zavodny (2005), Mishra (2007), Ibarraran and Lubotsky (2007), Fernandez-Huertas (2010), Caponi (2006) and McKenzie and Rapoport (2010).

Deininger and Feder (2009), and Galiani and Schargrodsky (2010b) for excellent reviews), whereas "the relationship between land tenure and off-farm labor market participation is under-researched, especially in rural areas of developing countries" (Deininger and Feder (2009):256). For urban areas, the evidence is mixed. Field (2007) finds a positive impact on labor supply outside the home among urban squatters in Peru, while Galiani and Schargrodsky (2010a) find no impact among urban squatters in Buenos Aires. Whether urban property rights have an impact on labor supply outside the home may depend on whether the labor supply was constrained prior to the change in property rights (Galiani and Schargrodsky (2010b)). For rural areas, Do and Iyer (2008) find a positive impact on off-farm labor supply among rural households in Vietnam, although it is ten times smaller than the impact identified by Field (2007).[2] To our knowledge, there is no evidence on the impact of land certification on migration, which is the natural extension of the study of non-farm labor participation. Since Mexican household members can now leave (and even rent out) their land without fear of being expropriated or fear of losing their inheritance, they may be able to migrate to higher-income work, which may imply urban areas or, in our case, the US.

The major added value of the paper is the identification strategy. Property rights are typically endogenous to household behavior (Besley and Ghatak (2010)). In order to tackle the corresponding identification challenge, we take the following steps. First, we consider a land certification program that provides a neat source of discontinuity in (de facto) property rights between certified and non-certified communities. Second, we use survey data on the same households prior to the program to control for all unobserved time-invariant differences between program and non-program areas that may be correlated with migration behavior.

Third, we control for unobserved time-varying differences between program and non-program areas, which may still be correlated with migration behavior, by using an additional control group (non-eligible households) and employing a DDD strategy.[3] This identification strategy is what distinguishes the present paper from Mullan, Grosjean, and Kontoleon (2011)

[2] Field (2007) finds an increase equal to 3.04 working hours outside the home per week per working household member, while Do and Iyer (2008) find an increase equal to 0.36, almost ten times smaller. In the latter paper there is no descriptive statistic on labor supply before (and after) the program, so we cannot speculate on the extent to which the labor supply was constrained.

[3] See Field (2007) for a similar approach.

and de la Rupelle, Quheng, Shi, and Vendryes (2009), who look at rural-urban migration in China, and de Braw and Mueller (2009), who look at internal migration in Ethiopia. In contrast to them, we use a land certification program (and a DDD strategy) to identify the causal impact of land property rights on migration, rather than self-reported tenure security or land transferability.

The paper is structured as follows: Section 2 describes the certification program and land property rights in Mexico; Section 3 discusses the theory linking land property rights to household migration behavior; Section 4 presents the data, the identification strategy, and the regression specification; Section 5 presents the results; and Section 6 concludes the paper.

2 Context: *Procede* in Mexican ejidos

Following the 1911 revolution, the Mexican government established that groups of farmers could free of charge receive non-transferable land in usufruct.[4] The ejido is the agrarian institution that is endowed with such land and which is generated with this application (Quesnel (2003)). The ejidatarios are the farmers who applied for such land. They could decide whether to divide part or all of the land into individual plots.[5] Each of them received one individual plot and access to the common land. Individual plots were used mainly for rainfed agriculture, while common land was used mainly for cattle and livestock grazing (Procuraduaria Agraria (2010)).

Throughout the decades ejidos arrived to include an estimated 3.2 million ejidatarios in about 30,000 ejidos and to constitute 56 percent of the national land usable for agriculture (World Bank (1999)).[6] Ejidos became characterized by levels of capital endowment significantly lower than in the private sector (World Bank (2001)) and by extreme poverty (Velez (1995)).

[4] Article 27 of Estados Unidos Mexicanos (1917).

[5] Details can be found in Estados Unidos Mexicanos (1971). See articles 130, 134 and 135.

[6] The remaining land used for agriculture is private property and is not considered in our empirical application.

The 1992 Agrarian Law grants ejidatarios full property rights to their urban plots, the rights to sell (exclusively to members of the same ejido) and rent out their individual plots,[7] and the right to use the common land, but not to transfer it.[8]

The law confirms the use rights on all plot types and introduces the transfer rights on urban and individual plots. In addition, it introduces the rights to use wage labor and to leave the individual plots fallow for more than two years.[9] The limits to the right to sell imply the virtual impossibility to collateralizing land to obtain credit.[10]

At the end of 1993 the government launched a massive certification program, called *Procede*. As part of the program, ejidatarios' rights over land were documented with certificates issued by the National Agrarian Registry (RAN).

Certificates for individual plots (*certificado parcelarios*) included the name of the ejidatario, the size and position of the plot, and the list of bordering neighbors. The certificates replaced the old certificates (*certificado de derechos agrarios*), which included only the name, the ejido affiliation, and the way of acquisition of the plot (Del Rey Poveda (2005):162,166). Certificates of access to common land reported the ejidatario's name and the proportion of the common land he/she had the right to use.

Procede aimed to provide certificates to all ejidatarios, i.e., they were all *eligible* for the program. Non-eligible landed households in the ejidos were households with no formal rights to land, either because they had no blood ties with the farmers in the ejido or because they had blood ties but the household head did not inherit the land. This group came to possess land through occupation of empty plots or acquisition through black markets, and arrived to constitute 37.2 percent of agrarian subjects (World Bank (2001):13-14). They did have the

[7] See articles 68, 79 and 80 of Estados Unidos Mexicanos (1992).

[8] Only the ejido Assembly, in case of majority of votes, has the right to transfer the common land. Such right is limited to the common land as a whole and to companies external to the ejido (art.75) and does not seem to have been used in practice.

[9] Details of ejidatarios' rights can be found in Estados Unidos Mexicanos (1971). For rights on urban plots, see article 93. For rights on individual plots, see articles 52, 55, 77 and 85. Possible exceptions are listed in article 76. For rights on common land, see article 67.

[10] A plot can be used as collateral only with credit institutions that already have commercial relationships with the ejido, and, in case of default, the credit institutions can seize the plot only for the amount of time necessary to get the money (Art. 46). So, we do not expect certificates to have increased access to credit. Acquisition of full property rights (*dominio pleno*) requires an additional deliberation of the Assembly and an individual application of the ejidatario to the RAN (Art.81-82). In practice, very few Assemblies seem to have done so. Only 6/248 ejidos in our sample have adopted dominio pleno.

right to buy one urban plot (but not to trade it further), which made them eligible for the housing title, but no right to individual or common land, making them *non-eligible* for the certificates.

Rather than simply imposing the program on the communities, government officials visited and informed them. Adoption required the consent of a large majority of ejidatarios.[11] The issuance of certificates was relatively successful. *Procede* resulted in the issuance of "certificates to more than 3 million households" (World Bank (2001)).

The certification constituted a *de facto* change in land property rights (as opposed to a *de jure* change), because, rather than providing rights, it improved ejidatarios' ability to take advantage of their formal property rights.[12]

3 Theoretical framework

How can we expect better land property rights to affect migration? The seminal paper by Besley (1995) and the recent survey by Besley and Ghatak (2010) provide a simple framework which, applied to our context, suggests that better property rights unambiguously increase investments via less fear of expropriation (by the state and by other households) and gains from trade.[13] International migration is a highly remunerative type of off-farm labor supply. A simple extension of this argument to include off-farm labor supply predicts a decrease in off-farm labor supply if investments are labor-intensive (e.g., manure, land clearing, and adoption of labor-intensive crops) and an increase if investments are capital intensive (e.g., machinery,

[11] Estados Unidos Mexicanos (1992) describes the adoption procedure in detail. The beginning of the certification program required the head of the village (Comisario Ejidal) to call for the "Information and Consent Assembly". This assembly required the presence of the simple majority of ejidatarios (first call), or any number of them (successive calls), to be valid (art.26). It also required the approval of the simple majority of them to allow officials to map the ejido (art.27). After the measurement took place, the head of the village had to call for the "Delimitation, Assignment and Entitlement Assembly". This assembly required the three fourth of ejidatarios (first call), or its simple majority (successive calls), to be valid (art.26). It also required the approval of two thirds of them (art.27) for the map to be sent to the cadastre (RAN) to be registered. The program terminated when the ejidatarios received the certificates from the cadastre.

[12] Differently from the certification program, the 1992 Agrarian Law applied immediately to all ejidatarios, independently from the possession of the new certificates. Article 4 Transitorios states that ejidatarios in non-program areas maintain their status and can take advantage of the provisions of the 1992 Agrarian Law.

[13] A third channel, collateralizability of land, does not seem to be at work in our context (section 2).

fertilizer, and cattle).[14]

In this paper we formalize an additional mechanism recently suggested by Galiani and Schargrodsky (2010a): the fear of expropriation from within the family.[15] Before the 1992 Agrarian Law, ejidatarios transmitted rights over land only through inheritance. The heir had to be unique, but the ejidatario could choose him/her by stating an order of preference. If he did not do so, the law gave priority to the wife/husband and then to the children, where the order among the latter was left unspecified. If the inheritance went to the children, the ejido assembly intervened to determine the heir.[16] When doing so, the assembly took into account the ability and willingness of the (potential) heir(s) to take charge of the inheritance (Del Rey Poveda (2005):163,173).

This encouraged strategic behavior by the potential heirs (Del Rey Poveda (2005):182). Signaling an ability to take charge of the land and a willingness to remain in the ejido constituted an incentive against migration, since leaving was a clear signal of weak attachment to the land (Del Rey Poveda (2005):170,184). This is consistent with anecdotal evidence from Western Mexico:

> The child who looks after the parents until their death develops certain rights to the property. This may sometimes lead to awkward situations among brothers and sisters who do not want one sibling to look after their parents too much and in this way create claims to the land. (..) Alternatively, a son who has migrated to the United States and declares that he does not intend to come back, may be replaced as heir by a son in the village. (Nuijten (2003):486).

The 1992 Agrarian Law maintains the same inheritance rule with one caveat: potential heirs have three months to find an agreement or the Agrarian Tribunal (rather than the ejido

[14] This channel refers to migration as a self-funding strategy, which is supported by evidence of a positive impact of migration (or remittances) on agricultural technology (Mendola (2008)), household investments (Yang (2008)), and entrepeneurship (Woodruff and Zenteno (2007)). See also de Janvry, Gordillo, and Sadoulet (1997) for a description of the migration-subsistence strategy of Mexican farmers.

[15] "The lack of titles may also impede the division of wealth among family members, forcing claimants to live together to enjoy and retain usufructuary rights" (Galiani and Schargrodsky 2010:708).

[16] See articles 81 and 82 of Estados Unidos Mexicanos (1971).

assembly) will proceed to sell the land within the ejido and split the revenue among the children in equal shares (Del Rey Poveda (2005):163; Riveros Fragoso (2005):44).[17]

There is strong evidence that resorting to the Agrarian Tribunal to settle disputes over land inheritance was a feasible option. The Agrarian Tribunal dealt with more than 104,000 cases concerning land inheritance out of a total of 315,000 during the period 1992-2005 (Morales Jurado and Colin Salgado (2006):229).[18] Land inheritance is by far the primary issue dealt with in terms of number of cases. Even more interestingly, data from the Procuraduria Agraria show that the number of land inheritance law cases has increased dramatically in ejidos that implemented the program (Figure 1).

Thus, certification improves access to courts; potential heirs can now contest land inheritance through outright negotiation in the shadow of the Agrarian Tribunal and no longer have to be present in the ejido. A simple way to capture the influence of better property rights on off-farm labor supply via the land inheritance mechanism is to consider a two-period extension of the basic agricultural model (Singh, Squire, and Strauss (1986)),[19] where the decision maker is the single household member rather than the household as a whole.

Household member i allocates his/her labor supply ($\bar{T}$) to in-farm (T_{if}) and off-farm (T_{io}) activities.[20] Let $Y(T_f, L)$ denote the agricultural production given labor supply T_f and land input L. The function $Y : \mathbb{R}^2_+ \to \mathbb{R}$ denotes the agricultural technology. Assume that

Assumption 1. Y is continuous, twice differentiable, increasing and concave in each argument with $\lim_{T_f \to 0} Y_1(T_f, L) = Y_1(0, L) = \infty$.

[17] See articles 17 and 18 of Estados Unidos Mexicanos (1992).

[18] The importance of the definition of the heirs is confirmed by the HEREDA program (Procuraduaria Agraria (2007):169). The HEREDA program started in 2001 and aims at letting all household heads write down a will.

[19] See Chiappori and Donni (2009) for a review of the literature on non-unitary household models. See Browning et al. (2006) for a comparison between unitary and non-unitary household models. Within the migration literature, see Rapoport and Doquier (2006) for a review of the literature on migration and remittances using non-unitary household models.

[20] Off-farm activities include local off-farm activities, domestic migration, and international migration. As long as temporary and return migration are relatively common and the time horizon is medium rather than short, international migration may be considered a continuous choice.

We abstract from the presence of leisure to keep the model mathematically tractable. We also abstract from any distinction between in-farm (productive) labor and guard (unproductive) labor. This is motivated by the fact that: i) guarding is this case is just a signal and does not require specific time or effort; ii) any distinction would be unobservable at the empirical level (in a rural context).

In the first period all household members pool their in-farm labor supplies $\left(T_f = \sum_i T_{if}\right)$. In return, each of them receives an equal share of the agricultural product: $\frac{1}{N}Y(T_f, L)$. In the second period, only the member who captured the land can devote in-farm labor supply to it $(T_f = T_{if})$. In return, he/she received the entire agricultural product: $Y(T_f, L)$. Let w denote the return from each-unit of labor supply devoted to off-farm activities[21].

We assume that household members can influence future land allocation by working in the in-farm activity. The idea is that working the land strengthens the claims over it[22]. On the other hand, an eventual dispute could be settled through a court, be it an Agrarian Tribunal or a less formal local village council. The ability of courts to intervene and settle the dispute increases with land property rights (θ). Weak property rights over land leave room for expropriation from other households (E).

Define the winning probability of member i as a function of own in-farm labor-supply (T_{if}), others' in-farm labor supplies (T_{kf}, with $k \neq i$), external labor supply (T_E) and land property rights (θ) in the following way:

$$p^i = \begin{cases} p\left(\frac{f(T_{if1})}{f(T_{if1})+\sum\limits_{k\neq i} f(T_{kf1})+f(T_E)}, \theta\right) & if \quad f(T_{if1}) + \sum\limits_{k\neq i} f(T_{kf1}) + f(T_E) > 0 \\ p\left(\frac{1}{N}, \theta\right) & otherwise \end{cases},$$

where $p_1 > 0, p_{11} < 0, p_2 > 0, p_{22} < 0$, and $p_{12} < 0$. The first argument corresponds to a rather general contest success function, where $f' > 0$ and $f'' < 0$ (see Skaperdas (1996) for an axiomatization and Garfinkel and Skaperdas (2007) for a review of the literature). The key assumption is that labor supply and property rights are substitutes in the land dispute.

[21] Clearly, when we consider migration w is the return net of all variable costs. Such costs are expected monetary and non-monetary, where the non-monetary component can be substantial (Hanson (2010)). In case of international migration there is also a substantial fixed costs. This is trivial to add to the model and it will be considered in the empirical analysis.

[22] Since we don't model heterogeneity across members of the same households, if they do not contest the land their payoff is homogeneous across members. This could be interpreted either as equal probability of inherit the land or equal division of the land inheritance. The latter could take place either directly by division of the land, or indirectly through assignment of the land to the heir and monetary compensation to the others.

It would be possible to include some degree of heterogeneity across members through the contest success function. This could account for specific inheritance rules like primogeniture. However, this would not alter the qualitative prediction of the model.

This assumption captures the idea household members' access to courts is increasing with the available documentation.

The timing is the following:

- all household members choose simultaneously their labor supply allocation (T_{if1}, T_{io1});
- nature chooses the heir with probabilities p^i;
- the heir allocates his/her labor supply (T_{if2}, T_{io2}).

The generic member's decision problem in the first period is:

$$\max_{T_{if1}, T_{io1}} \frac{1}{N} Y(T_{if1} + \sum_{k \neq i} T_{kf1}, L) + wT_{io1} + \delta \left\{ p^i \left[Y(T_{if2}, L) + wT_{io2} \right] + \left(1 - p^i\right) w\bar{T} \right\}$$

$$s.t. \begin{cases} T_{if1} + T_{io1} = \bar{T} \\ T_{if1}, T_{io1} \geq 0 \end{cases}$$

In case i becomes the heir, his/her decision problem in the second period will be:

$$\max_{T_{if2}, T_{io2}} \{Y(T_{if2}, L) + wT_{io2}\} \text{ s.t. } \begin{cases} T_{if2} + T_{io2} = \bar{T} \\ T_{if2}, T_{io2} \geq 0 \end{cases}$$

It turns out (see the Appendix for a detailed analysis) that whoever captures the land finds worthwhile to devote some labor to it. This makes competition for the land asset salient in the first period, which is when the strategic interaction takes place. In equilibrium all members devote the same amount of in-farm labor-supply and this amount is positive.

Concerning the relationship between (first-period) labor-supply and land property rights, the following result applies:

Proposition 1 *Suppose that assumption 1 holds. Then household members' in-farm labor-supply is decreasing in land property rights, while household members' migration is increasing* *in land property rights*[23]*:*

$$\frac{dT^*_{fi1}}{d\theta} < 0 \text{ } and \text{ } \frac{dT^*_{oi1}}{d\theta} > 0.$$

Since the proposition applies to each household member, it applies implicitly to the household as a whole: $\frac{dT^*_{f1}}{d\theta} < 0$ and $\frac{dT^*_{o1}}{d\theta} > 0$.

4 Data and estimation method

4.1 Data

We consider the 1994 and 1997 ejido surveys. The 1994 survey was carried out by the Mexican Ministry of Agrarian Reform (Segreteria de Reforma Agraria, SRA) in collaboration with University of California Berkeley and is designed to be nationally representative of all ejidos (and communities) in Mexico.[24] The 1997 survey was carried out by the Ministry of Agrarian Reform with the World Bank following the same survey design as in 1994. The surveys provide information on 1,286 panel households.[25]

The surveys provide detailed information on household members' demographic characteristics, past migration experiences, current migration experiences of children of the household head living outside the house, use of land, equipment, and ejido characteristics.[26]

[23] If the members' equilibrium in-farm labor supply happens to be a corner solution ($T^*_{if1} = \bar{T}\ \forall i$), then in-farm labor (migration) is weakly decreasing (increasing) in land property rights.

[24] The survey is representative at the state level. Ejidos were selected from each state except Chiapas, where conflict prevented fieldwork. Details can be found in de Janvry, Gordillo, and Sadoulet (1997).

[25] The attrition rate was only 4.0%. See World Bank (1999): Annex 2 for details. The program started between 1993 and 1994, i.e., only a few months before the 1994 survey, which was conducted during the summer. We exclude 14 households as they belong to ejidos with missing information regarding the program, 108 households as they belong to ejidos that completed the program before the 1994 survey, 15 households because they are private landowners, 113 households due to unclear status (to be specified later), and 110 households because they belong to communities instead of ejidos. The final sample has 926 households in 221 ejidos.

[26] These data have been used by several other authors for a variety of purposes: ejido reforms (World Bank (1999), World Bank (2001), Munoz-Pina, De Janvry, and Sadoulet (2003), migration (Winters, de Janvry, and Sadoulet (2001); Davis and Winters (2001)), off-farm activities (de Janvry and Sadoulet (2001)) and cash transfer programs (Sadoulet, Janvry, and Davis (2001)).

4.2 Migration to the United States

Mexicans started migrating to the US from rural areas following the construction of railroads in the early 20th century and the Bracero program from 1942 to 1964 (Hanson 2006). De Janvry, Gordillo, and Sadoulet (1997) show that the variation in migration experience among ejidatarios' cohorts is consistent with them having been part of this migration flow. Out-migration is historically high in the northern and central regions. These regions also constitute the primary location of ejidos; our final sample of ejido households is located primarily in the central (29.48%) and northern (22.57%) regions, followed by the Gulf (17.28%), south Pacific (16.95) and north Pacific (13.71%) areas. The distribution of ejido households across Mexican states is positively but not perfectly correlated with the 1994 population distribution for the entire Mexico (the state-level correlation is 0.44). In turn, state migration rates are positively correlated with the distribution of ejido households (0.30) but not with the population distribution (-0.02).[27]

In order to identify migrant households we construct a binary indicator taking the value one if any household member who is currently living at home has been in the US within the previous three years or if any child of the household head currently lives in the US. Migrant households amount to 15 percent in 1994 and 29 percent in 1997. The average number of migrants per household is 0.3 in 1994 and 0.72 in 1997. These migration rates are consistent with Winters, de Janvry, and Sadoulet (2001) for 1994 and with Davis and Winters (2001) for 1997. The increase in the number of migrants from 1994 to 1997 (0.420) corresponds to about 1,384,281 additional migrants (both temporary and permanent).[28] U.S. Immigration and Naturalization Service (2003) provides some yearly estimates of the number of illegal Mexicans who entered the US during the period 1990-1999; the number of additional migrants for the period 1994-1997 is 1,873,000 illegal entrants. These estimates rely on assumptions

[27] Conteo de Poblacion y Vivienda (1995). Own tabulations. Migration is defined as the share of the population that migrated to the United States within the previous five years.

[28] The number of additional migrants is obtained by multiplying the number of ejidos (26,796, according to World Bank 2001) with the average number of landed households per ejido (123) and the increase in the number of migrants per landed household (0.420). Using the estimates in Winters and Davis (2001), one obtains 875,184 additional migrants, perhaps because they include "comunidades", which typically have low migration rates.

of under-counting and should be used cautiously. According to Hanson (2006), the true flow could be 15 percent higher than the estimate reported by INS, i.e., 2,153,950 entrants. During the same period, the number of legal Mexican migrants was 511,883 (U.S. Immigration and Naturalization Service (1999)). Hence, the total number of migrants is between 2,384,883 and 2,665,883. Based on these estimates, the 1994-1997 increase in the number of migrants from Mexican ejidos corresponds to 52-58 percent of the number of Mexicans who entered the US. This is consistent with migration stemming primarily from rural areas and ejido households constituting a large fraction of the rural population.[29]

4.3 Identification strategy

In this paper we exploit both the timing of the certification program and heterogeneity in farmers' status within ejidos to identify the impact of the program on household migration behavior. The 1997 ejido survey contains detailed information on the implementation of the program. Ejidos that completed the program before the 1997 survey are termed "program areas," whereas those that did not are termed "non-program areas." Households in non-program areas constitute our first control group. Ejidatarios in program areas benefit from the program as they receive the certificate for their houses and their individual plots as well as for access to common land.[30]

Program timing may be far from randomly allocated: government officials may have implemented the program according to ease of entry; the decision to implement the program by the ejido assembly may have suffered from collective action problems and from the resolution of internal land conflicts. Table 1 shows the self-reported explanations for the decision to implement or not implement the program. As can be seen, the primary reason to implement the program was tenure security (88.3%), followed by willingness to solve border issues

[29] According to de Janvry (1995) ejidos include 70 percent of all Mexican farmers.

[30] In the 1997 ejido survey, 13% of ejidatario households in program areas report no *Procede* certificate for their individual plots. An additional 9% report to have receive *Procede* certificates for some but not all their plots. The (unobserved) reasons could be the following. First, some of the certificates might have not arrived yet. This is consistent with relatively low certification rates in ejidos certified in 1997 and in ejidos where the date of reception of the certificates is missing. Second, households may own land in ejidos, different from the one they live in, which have not been certified yet. Partial and delayed certification makes the estimation of the LATE of the certificates problematic.

(29.7%); the primary reason not to implement the program was lack of information (30.4%), tax avoidance (15.9%), and border issues (15.9%). Overall, these explanations are certainly interesting, yet the only surprising feature is the small role played by land market motives. We will make use of some of this information later in the analysis.

In Table 2 we compare some observable ejido characteristics across program and non-program areas prior to the program (Columns 1-3). Program areas have a higher percentage of parceled land relative to common land, less ejidatarios, a more equal distribution of parceled land, better infrastructure (access to paved road, electricity, drinking water and drainage, existence of an assembly hall), and fewer boundary problems. The differences suggest that the program may have been directed to smaller and wealthier ejidos first, which is consistent with World Bank (1999) and World Bank (2001).

Non-random program timing may be problematic if the determinants of program implementation are correlated with household migration behavior. In order to correct for this bias, we could control for ejido characteristics that we found to be correlated with program implementation (selection-on-observables). However, there would be no way for us to be sure of having included all relevant determinants.[31]

In order to improve our identification strategy, we make use of non-eligible households as an additional control group and compare the difference in migration behavior between eligible and non-eligible households in program areas with the difference between eligible and non-eligible households in non-program areas. Let M_i be an indicator for the migration behavior of household i and let P and E indicate program areas and eligible status, respectively. Our baseline comparison is:

$$\{E[M_i|P=1,E=1]-E[M_i|P=1,E=0]\}-\{E[M_i|P=0,E=1]-E[M_i|P=0,E=0]\}.$$

Let $M_i(P,E)$ denote potential outcomes and assume that the program is randomly allo-

[31]Two potential confounding factors are the pre-NAFTA subsidies and migration networks. Entry into NAFTA led to the removal of subsidies to agriculture and, possibly, to out-migration (de Janvry and Sadoulet (2001), Sadoulet, Janvry, and Davis (2001)). This may bias our estimates if pre-NAFTA subsidies differed across program and non-program areas. The same is true for community migration networks (Winters, de Janvry, and Sadoulet (2001), Munshi (2003)).

cated across eligible and non-eligible households:

$$E[M_i(0,1)|P=0,E=1]-E[M_i(0,0)|P=0,E=0]=$$
$$=E[M_i(0,1)|P=1,E=1]-E[M_i(0,0)|P=1,E=0].$$

Then we can re-write (see Appendix) the baseline comparison as:

$$E[M_i(1,1)-M_i(0,1)|P=1,E=1]-E[M_i(1,0)-M_i(0,0)|P=1,E=0].$$

This expression corresponds to the *mean effect of the program on eligible relative to non-eligible households*. Since one of the control groups (non-eligible households in program areas) gets partial access to the program, the potential outcomes within the second part of the expression do not cancel out and the estimator corresponds to a downward biased estimator of the *mean effect of the program on eligible households* (Heckman, Lalonde, and Smith (1999)).[32] Non-eligible households in program areas receive the certificates for their housing plots; they do not receive the certificates for their individual plots unless the ejido assembly recognizes them in their status of possessors (which happens 66 percent of the times); they do not receive the certificates of access to common land unless the ejido assembly upgrades them to ejidatario status (which happens, on average, 34 percent of the cases).[33]

In order to identify eligible and non-eligible households, we make use of pre-program (1994) data on possession of an ejido certificate. Households with a pre-program ejido certificate are termed "eligible," whereas those without are termed "non-eligible."[34] An informal check of

[32] The econometric issue is very similar to control group members having access to a substitute program (Heckman, Hohmann, Smith, and Khoo (2000)) and to a measurement error in "eligibility" status among comparison group members (Heckman, Lalonde, and Smith (1999), Heckman and Robb Jr (1985)). It is not clear whether both mean effects are Intent-To-Treat (ITT) effects or not. For example, in Banerjee, Duflo, Glennerster, and Kinnan (2010), part of control group members access the program and the authors still present their estimator as an ITT.

[33] This share is the outcome of the following back-of-the-envelope exercise: in 1994 there were 87 eligible households in program areas (Table 2); the ratio ineligible-eligible households in program areas in our sample is 0.57, i.e., an average of 50 ineligible households in program areas; from 1994 to 1997 the number of eligible households in program areas increased from 87 to 104, which corresponds to an upgrading of 34 percent of ineligible households.

[34] According to Estados Unidos Mexicanos (1971) (Art. 69) and to Del Rey Poveda (2005):166, ejidatarios' rights are acknowledged by certification (*certificado de derechos agrarios*). Indeed, these certificates constitute

the quasi-random assignment of the program across eligible and non-eligible households is to compare observable characteristics of eligible and non-eligible households across program and non-program areas *prior* to the program. The results (Table 2) show a lack of significant differences across groups (Column 9) in migration rates, household demographics, dwelling characteristics, assets, and land transactions. Besides, even the comparison of each group of households across program and non-program areas (Columns 3-5, 6-8) shows very little differences.[35] Households' pre-program tenure security is unobserved, but there are strong theoretical reasons to expect tenure security to be correlated with the intensity of land transactions (Besley (1995), Besley and Ghatak (2010), and Deininger and Feder (2009)). Table 2 shows that land transactions were relatively widespread prior to the program, and that their intensity does not differ across groups. This is consistent with case studies (Nuijten (2003)) suggesting that informal tenure security was relatively strong and supported widespread black markets.[36]

The 1997 ejido survey also includes information on the date of completion of the program. This will allow us to separate program areas into early (1994-1995) and late (1996-1997) program areas. This differentiation captures the fact that households in early program areas had more time to adjust their migration behavior. It may therefore also be appropriate to compare eligible and non-eligible households across early and late program areas (Table A2 in the online appendix). Notwithstanding the limited sample size, there are remarkably few differences between eligible and non-eligible households across early and late program areas (Column 8).

By using non-eligible households as an additional control group, we control for all differences across program and non-program areas shared by the two groups. Still, it could be that migration behavior differs between eligible and non-eligible households across program and non-program areas due to factors other than the certification program.

One way to relax this identification assumption is to control for household-level charac-

the basis for the delivery of the new certificates (Art.4 Transitorios, Estados Unidos Mexicanos (1992)).

[35] Table A1 (online appendix) confirms the comparability of the two groups across program and non-program areas with 1997 data.

[36] In fact, pre-1992 land transactions were illegal but widely accepted within ejidos (Yates (1981):181, and NACLA (1976):18, cited in Heath (1990):34).

teristics, which we select based on the migration literature. Descriptive statistics comparing migrant and non-migrant households (not reported) show that migrant households are bigger, associated with a greater number of siblings of the household head abroad,[37] less likely to be indigenous, and associated with greater land assets and better dwelling characteristics. On the other hand, their household heads are older and less educated (but equally literate). Average schooling is similar.[38]

Another way to relax our identification assumption is to exploit the time-series dimension of our dataset. By doing so, the identification assumption is that the difference in migration behavior between ejidatarios and non-ejidatarios across program and non-program areas does not vary over time due to factors other than the certification program. Thus, we allow for a difference in migration behavior, but it must be constant over time.

4.4 Regression specification

The model presented in Section 3 predicts that an increase in land property rights causes a decrease in in-farm labor supply and an increase in off-farm labor supply. The prediction is valid both at the individual and household level. In this section we will test the prediction at the household level. Since the household surveys are rich in questions on household members' migration experiences but not on in-farm labor supply, we will focus on the former. The outcome of interest is household migration status (see Sub-section 4.2). As a robustness check, we will also report the results for the number and for the share of migrant members.

We estimate 1997 household migration status with the following Linear Probability Model (LPM):

$$y_{ik} = \eta_1 + \alpha_1 w_i + \beta_1 (w_i * e_{ik}) + \gamma_1 e_{ik} + \Gamma'_{11} Z_{ik} + \Gamma'_{12} (Z_{ik} * e_{ik}) + \Gamma'_{13} X_i + \varepsilon_{1ik}, \qquad (1)$$

[37] The number of siblings of the household head abroad is a proxy for the strength of the household migration networks (Winters, de Janvry, and Sadoulet (2001)).

[38] The absence of selection in terms of education is surprising with respect to the literature on Mexican migration. However, note that the average level of education is very low in our sample (3-4 years of schooling), while Chiquiar and Hanson (2005) show that, in 1990, 73.9 percent Mexican residents had more than four years of education.

where $y_{ik} \in \{0, 1\}$ is the migration status of household k in ejido i, $w_i \in \{0, 1\}$ indicates whether ejido i completed the program before the 1997 survey, $e_{ik} \in \{0, 1\}$ indicates whether household k in ejido i is eligible, X_i is the vector of ejido-level controls, Z_{ik} is the vector of household-level controls, and ε_{1ik} is the error term clustered at the ejido level. We will also estimate the 1997 household migration status using a Logit model[39]. Equation (1) then corresponds to the latent variable specification. The household-level controls (Z_{ik}) are the following: household composition (age of the household head, number of adult members, fraction of females among adult members, average literacy[40], average schooling of adult members[41]), migration assets (number of siblings the household head abroad)[42], and land assets (land used in 1994). The ejido-level controls (X_i) are the following: land (ejido area in logarithm, share of common land with respect to common and parceled land), population composition (dummy for indigenous ejidos, membership to ejido union), and infrastructure (access to paved road).

The identification of the impact of *Procede* on eligible households (β_1) in (1) requires that there is no difference in migration behavior between eligible and non-eligible households across program and non-program areas driven by factors other than the program or the set of controls we include. This specification lets us control for all unobserved differences across program and non-program areas common to both eligible and non-eligible households (α_1), like distance from the border (which affects the cost of migration), historical community networks (which affect both the cost of migration and its expected return), and varying implementation of the program (due for example to administrative capacity of the Procuraduria Agraria across areas).

To address the possibility that the identification assumption does not hold, we exploit the time dimension of our dataset and estimate household migration status according to the following Pooled Linear Probability Model:[43]

[39] The marginal effect of the interaction term is computed according to Norton, Wang, and Ai (2004).

[40] This information is available for members currently living at home only.

[41] Adult household members are at least 15 years old.

[42] Notice that the siblings of the hosuehold head may have been part of the household before migrating. Therefore, our measure of household migration assets in 1997 may be partly endogenous to the program. In order to avoid this possibility, we consider its pre-program (1994) value.

[43] Again, we will also estimate household migration status using a Logit model (following Cornelissen and Sonderhof (2009) to compute the marginal effect associated with the triple interaction term).

$$
\begin{aligned}
y_{ikt} &= \alpha_{21} w_i + \alpha_{22}\left(w_i * 1997\right) + \gamma_{21} e_{ik} + \gamma_{22}\left(e_{ik} * 1997\right) + \gamma_{23} 1997 + \\
&\quad + \beta_{21}\left(w_i * e_{ik}\right) + \beta_{22}\left(w_i * 1997 * e_{ik}\right) + \Gamma_{21}' Z_{ik} + \Gamma_{22}'\left(Z_{ikt} * e_{ik}\right) + \varepsilon_{2ikt},
\end{aligned} \tag{2}
$$

where y_{ikt} is the migration status of household k in ejido i at time t, w_i is the dummy for ejidos that received certificates in 1997, and e_{ik} is the dummy for eligible households. The identification of the impact of *Procede* on eligible households (β_{22}) requires that the difference in migration behavior between eligible and non-eligible households across program and non-program areas, due to factors other than the program and the included controls, is constant over time. This assumption is weaker than the previous one, because now we control also for time-invariant unobserved differences between eligible and non-eligible households across program and non-program areas (β_{21}).

5 Results

5.1 Impact of *Procede* on migration

Table 4 shows the results associated with the cross-section specification (1). Without controlling for any background characteristics, the coefficient estimate associated with eligible households in program areas is positive and large (0.115), but not significant at conventional levels. We then control for background characteristics (Column 2): the coefficient is now larger (0.127) and marginally significant. The marginal effect associated with a Logit model (Column 3) has similar magnitude (0.119) and is also marginally significant. The result is robust to the use of alternative dependent variables, such as the number of migrants (Column 4) and the ratio of migrants to adult household members (Column 5).

The direction, magnitude, and significance of the coefficients associated with the control variables are quite consistent with basic economic theory; i.e., the opportunity cost of migration decreases with household size if agriculture is characterized by decreasing marginal returns (each additional adult increases the likelihood of migrant status by 3 percent), and cultural barriers and geographical distance from the US are associated with less migration (the coefficient associated with indigenous ejidos is negative in all specifications).

In order to find out the seriousness of the concern for endogenous selection into the program we restrict the sample to non-eligible households who did not receive any certificate and estimate a difference-in-difference model comparing program and non-program areas before and after the introduction of the program. Table 5, Panel A, shows the results: the coefficient associated with non-eligible households in program areas is negative, small, and insignificant (between -0.035 and -0.062).

Table 5, Panel B, shows the results associated with the panel specification (2). The coefficient estimate associated with eligible households in program areas is positive, large, and significant or marginally significant in all specifications (Columns 1-8). Since households in early program areas (1994-1995) had more time to adjust their migration behavior than households in late program areas (1996-1997), we re-estimate some of the specifications using program timing, which takes the value 1 for late program areas and the value 2 for early program areas (Columns 8-10). The coefficient estimate is positive and significant, and its magnitude is consistent with the baseline estimates. Note that the magnitude, which ranges from 0.112 to 0.129, is remarkably similar to the one associated with the cross-section specification, which suggests the absence of any unobserved time-invariant difference in migration behavior between eligible and non-eligible households across program and non-program areas.[44] The coefficient estimates associated with non-eligible households in program areas (program*1997) and eligible households in non-program areas (eligible*1997) are much smaller and generally insignificant, which is also reassuring[45].

A coefficient estimate of 0.12 is very large. It constitutes an increase in migration rates

[44] As a robustness check, we re-estimate specification (2) controlling for non-land household assets that had shown some differences across groups in Table 3. Since they may be affected by the program, we include pre-program assets in levels and interacted with the time dummy. Table A3 shows the results: the coefficient of interest is robust to these additional controls (0.112-0.118), although we lose some precision in some of the specifications.

[45] We also estimated a DD specification with sample restricted according to eligibility status. Table A4 shows the results for eligible households (Panel A), ineligible households (Panel B) and without distinction in terms of eligibility (Panel C).

of 80 percent relative to the 1994 average migration status (0.15) and 85.7 percent relative to the 1994-1997 time trend (0.14). Since eligible households in program areas are 32.2 percent of the entire sample, our coefficient estimate explains 27.6 percent of the overall 1994-1997 increase in migration. The land certification program appears to have had a profound impact on ejidatarios' migration behavior. In terms of number of migrants, our coefficient estimates correspond to 429,238 additional migrants.[46] As discussed in Sub-section 4.2, the number of migrants from Mexican ejidos during the period 1994-1997 equaled 1,384,281 people, while the number of Mexican migrants ranged between 2,384,883 and 2,665,883 people. Hence, the coefficient estimates explain 31 percent of the increase in Mexico-US migration from the ejido sector and 16-18 percent of the entire Mexico-US migration.

This magnitude can be explained in terms of great initial tenure insecurity. However, it is also consistent with the coefficient capturing part of the legal changes introduced with Estados Unidos Mexicanos (1992) (see Section 2). This would be the case if, for example, eligible households in non-program areas were not aware of such legal changes or presumed that they were conditional on the certification. In this case the impact of the program would capture not just a de facto change in property rights, but also a de jure one.

We know that implementation of the program required the substantial resolution of border issues within eligible households and between eligible and non-eligible households. Thus, one may worry that our selection into the program may be affected not just by the eligible households, but also by non-eligible households. If so, our identification strategy would fail to control for unobservable characteristics that could, in principle, be correlated with household migration behavior. We therefore re-estimate specification (2) excluding all the households within ejidos that report to have implemented (or failed to implement) the program to solve border issues or conflicts between eligible and non-eligible households. Table A5 shows the results: the coefficient associated with eligible households in program areas is positive, large, and marginally significant in all specifications. The magnitude is similar (slightly higher) as previously found: it ranges from 0.134 to 0.155. Thus, we find no evidence that this particular

[46] This magnitude is the result of the following expression: 26,796 (ejidos, according to World Bank 2001) *111/211 (share of program areas) *87.01 (average number of eligible households) *0.350 (impact on the number of migrants).

selection mechanism drives our results.

Finally, note that our theoretical model generates a prediction that may be applied not just to international migration but also to domestic migration and off-farm labor within the village. So far our analysis has focused only on the first margin. There are two reasons for this. First, the impact on international migration is arguably the most interesting among the three. Second, the survey was designed with a particular focus on international migration, whereas the emphasis on off-farm labor was not as strong. As regards domestic migration, we know whether household members migrated to another state. However, it is not possible to tell whether they migrated to an urban area within the same state or remained in the same village. Regarding off-farm labor supply, it would be desirable to know the number of in-farm and off-farm labor hours (like in Field (2007) and Do and Iyer (2008)). To this end, we will have to rely on the information about the primary and secondary occupation of household members living at home. Specifically, we estimate the impact of the program on non-agricultural status, i.e., at least one member currently living at home works outside agriculture. Table A6 shows the results: the coefficient estimate of interest is negative and never significant, and its magnitude varies across specifications. Thus, we find no evidence of an impact on off-farm labor for members currently living at home. This could be driven by measurement error in the dependent variable or simply be due to international migration absorbing the entire impact of the program on off-farm labor.

A subtle negative general equilibrium effect of the program has to do with social cohesion within the community. Community cohesion implies non-monetary ties that prevent people from migrating abroad (Hanson (2010)). The program may have damaged such cohesion. This would not bias our parameter estimate of interest if both eligible and non-eligible households were affected in the same way, while it could bias the coefficient upwards if eligible households were affected more than non-eligible ones. Fortunately, our community and household questionnaire includes a question on the effects of the program on social cohesion (only for program areas), reading: "If the ejido implemented the program, how has the program affected social cohesion? (more, same, less)." The fact that social cohesion was not affected

(67.77%) or even increased (22.51%) and that these percentages are identical across eligible and non-eligible households is reassuring.

5.2 Do differences in migration behavior reflect anticipatory responses to the program?

One may wonder whether the certification process may have led households to postpone their migration decision rather than having increased the incentive to send one or more household members abroad. For example, it could be that household members feared being left out from the certification process and therefore waited for the certificate to reach the household before deciding to migrate. It could also be that household members abroad returned home just before the program started to ensure that they would not lose future assets, and then went abroad again. If this were the case, we would be confounding a short-term behavioral response to the program with a structural change in the households' migration strategy. In terms of tenure security, we would mistakenly take short-term tenure insecurity generated by the program itself for a permanent increase in tenure security.

In order to rule out this possibility, we make use of future timing in specification (1) using the 1994 household survey. If there is anticipatory behavior, then households in early program areas should migrate less than households in late program areas. Table 6, Panel A, shows that the coefficient estimates associated with this exercise are insignificant and very close to zero, regardless of whether we consider program relative to non-program areas (Columns 1-6) or soon-to-be-certified areas (certified August-December 1994) relative to all others (Columns 7-9), and whether we add controls, use a non-linear model or alternative dependent variables.[47]

Second, the 1997 community questionnaire identifies non-program areas that have initiated but not completed the program (henceforth *in-process* areas). In contrast to the 1994 soon-to-be-certified areas, we do not know when the 1997 in-process areas will complete the program or whether they will do so before the areas that have not yet started it. If this distinction between non-program areas runs along the lines of some unobserved characteristic other than

[47] The results are similar if we extend the time window for soon-to-be-certified areas to ten months (August 1994 - May 1995).

the timing of the program, then our previous identification assumption does not guarantee the correct identification of the impact of the program on in-process areas or the impact of the program on program areas. Keeping this caveat in mind, we estimate the panel specification. Table 6, Panel B, shows the results: the coefficient estimate associated with eligible households in in-process areas is negative, relatively small, and insignificant; the coefficient estimate associated with eligible households in program areas is generally consistent with the previous findings, although slightly smaller, and not always precisely estimated.

Overall, anticipation issues do not seem to explain the evidence gathered so far, although we cannot exclude that they did play a minor role.

5.3 Impact heterogeneity and the inheritance channel

Impact heterogeneity may be used to identify the channel(s) through which the property rights-migration relationship takes place.[48] In Section 3, we suggested the land inheritance mechanism, i.e., uncertain property rights keep landless family members home as they fear to lose their land inheritance in case of departure.

In order to test this mechanism, we divide households depending on whether the household head has written a will and re-estimate specification (2) for each sub-sample. The program should have a strong impact on households with no will, as it reduces the relatives' need to stay home to defend their informal property rights over the land inheritance (since the certificate allowed them access to the Agrarian Tribunal to solve any dispute). Yet, we expect the program to have little or no impact on households with a will, as the identity of the heir is known and there is less room for dispute. Any competing rationale (Section 3) would have difficulties explaining heterogeneity of the impact of land property rights across households with and without a will. Table 7 shows that, in support of the inheritance mechanism, the coefficient of interest is positive, large, and significant among households without a will (Column 3: 0.147), while it is small and insignificant among households with a will (Column

[48] In the working paper version we also explore the impact heterogeneity with respect to land assets.

2: 0.039).[49]

It is important to recognize that such evidence is not conclusive. We do not know why some households have a will and some do not. Del Rey Poveda (2005:185-186) argues that household heads may avoid writing a will to reduce their children's willingness to migrate. This concern does not seem very problematic, as it may work as an attenuation bias.

A more serious concern is whether the program led households to write a will. There is anecdotal evidence suggesting that, while implementing the program, officials suggested that households deposit a will (Del Rey Poveda (2005):179). If eligible household heads with low propensity to migrate wrote down a will following the program to a larger extent than non-eligible household heads did, then the coefficient estimate associated with households with a will (Column 2) is downward biased, while the coefficient estimate associated with households without a will (Column 3) is upward biased. Fortunately, this is not what our data suggest. The distribution of wills across households (in 1997)[50] is 25% and 34% respectively for non-eligible and eligible households in non-program areas, and 45% and 37% respectively for non-eligible and eligible households in program areas. Thus, it seems that the program led more non-eligible household heads to write a will than eligible ones, rather than the other way around. If the decision to write a will was somehow related to migration behavior, it would have to work like an attenuation bias. Nonetheless, we know too little about the determinants of the decision to write a will (and our data do not allow for much more than what we do here), and hence we interpret the evidence in Table 7 as an interesting correlation rather than as conclusive evidence.

In Table A7 we look at two other potential channels: land rental transactions (Panel A) and wage non-family labor (Panel B). In both cases the outcome is a binary variable indicating whether the household has been involved in a land transaction within the previous three years, and whether the household has hired non-family labor within the previous 24 months. In both cases, the coefficient estimate of interest is always small and never significant.[51] Another

[49] It is also consistent with a slightly different rationale (included in the model in Section 3), i.e., rather than attenuating the competition among potential heirs, land property rights attenuate the fear of expropriation by other community members.

[50] The information about households' will is only available for 1997.

[51] The results are the same if we consider the number of land rental transactions.

outcome it would be interesting to consider is land sales transactions, but they are too few in our sample to even try to estimate a model. Thus, we find no evidence supporting channels other land inheritance.

6 Conclusion

In this paper we ask whether there is a relationship between land property rights and international migration. We identify the impact of land property rights by making use of a country-wide certification program in Mexico ejido sector. Specifically, we exploit both the gradual introduction of the program and households' eligibility status.

Comparing eligible and ineligible households, we find that the program increased the likelihood of having one or more members abroad by 12 percent. The result is robust to the use of alternative econometric models and dependent variables. In terms of number of migrants, our coefficient estimates explain 31 percent of the 1994-1997 increase in Mexican migrants from ejido areas and 16-18 percent of the increase from the entire Mexico.

We also find some evidence that the impact of the program occurred through the land inheritance channel, initially suggested by Galiani and Schargrodsky (2010a). The land inheritance channel implies that household members refrain from migrating because they worry about losing their land inheritance. Better land property rights attenuate this problem, thus acting as a substitute for a well-defined land inheritance rule. Consistent with our model, the impact on migration is strongest in households where the landowner has not provided a will. It is difficult to reconcile this correlation with alternative rationales.

Evidence of a relationship between land property rights and international migration is interesting also for other reasons. Notwithstanding its recent increase, the level of global migration is rather low (3% of world population). This is at odds with a high cross-country wage differential and the cost of crossing borders illegally, which for at least some countries is non-prohibitive. Our analysis suggests that weak land property rights constitute a (typically unobserved) migration cost. This finding may help reconcile the puzzle.

Although the results are specific to Mexico, whose proximity to the US makes it the country with the largest stock of emigrants in the world, it would not be surprising to find similar effects for other countries as well, although possibly limited to internal migration. In 2009 the World Bank allocated about US$1.5 billion to 46 Land Administration Projects all over the world (Deininger and Bell (2010)). Many of them have emigrant to population ratios greater than Mexico (Azerbaijan, Bosnia and Herzegovina, Kyrgyz Republic, Macedonia, Nicaragua, Tajikistan and Ukraine).[52] It would be interesting to investigate whether the studied relationship holds for these countries as well.

[52] See World Bank (2011). All countries mentioned have emigrant to population ratios above 10 percent. The Philippines, which is also implementing a Land Administration project, has a ratio just below 10 percent.

References

BANERJEE, A., E. DUFLO, R. GLENNERSTER, AND C. KINNAN (2010): "The Miracle of Microfinance? Evidence from a Randomized Evaluation," .

BESLEY, T. (1995): "Property Rights and Investment Incentives: Theory and Evidence from Ghana," *Journal of Political Economy*, 103(5), 903.

BESLEY, T., AND M. GHATAK (2010): "Property Rights and Economic Development*," in *Handbook of Development Economics*, ed. by R. Dani, and R. Mark, vol. Volume 5, pp. 4525–4595. Elsevier.

CAPONI, V. (2006): "Heterogeneous Human Capital and Migration: Who Migrates from Mexico to the US?," Mimeo.

CHIQUIAR, D., AND G. HANSON (2005): "International Migration, Self Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 113(2), 239–281.

CORNELISSEN, T., AND K. SONDERHOF (2009): "Partial effects in probit and logit models with a triple dummy-variable interaction term," *Stata Journal*, 9(4), 571–583.

DAVIS, B., AND P. WINTERS (2001): "Gender, Networks and Mexico-US Migration," *Journal of Development Studies*, 38(2), 1.

DE BRAW, A., AND V. MUELLER (2009): "Do Limitations in Land Rights Transferability Influence Low Mobility Rates in Ethiopia?," .

DE JANVRY, A. (1995): *Reformas del sector agricola y el campesiando en Mexico*. San Joseť: International Fund for Agriculture and Development.

DE JANVRY, A., G. GORDILLO, AND E. SADOULET (1997): *Mexico's Second Agrarian Reform: Household and Community Responses, 1990-1994*.

DE JANVRY, A., AND E. SADOULET (2001): "Income Strategies Among Rural Households in Mexico: The Role of Off-farm Activities," *World Development*, 29(3), 467–480.

de la Rupelle, M., D. Quheng, L. Shi, and T. Vendryes (2009): "Land Rights Insecurity and Temporary Migration in Rural China," .

Deininger, K., and K. C. Bell (2010): "World Bank Policy on Land Governance in Support of the Millennium Development Goals," .

Deininger, K., and G. Feder (2009): "Land Registration, Governance, and Development: Evidence and Implications for Policy," *The World Bank Research Observer*, 24(2), 233–266.

Del Rey Poveda, L. A. (2005): "El nuevo marco de relaciones intergeneracionales en las familias ejidales: migracion y herencia en el sur de Veracruz," *Estudios Agrarios (Procuraduria Agraria)*, 28, 151–193.

Do, Q.-T., and L. Iyer (2008): "Land Titling and Rural Transition in Vietnam," *Economic Development & Cultural Change*, 56(3), 531–579.

Estados Unidos Mexicanos (1917): "Consitucion Politica de los Estados Unidos Mexicanos," SecretarÃŋa de Servicios Parlamentarios.

——— (1971): "Ley de la Reforma Agraria," .

——— (1992): "Ley Agraria. Nueva Ley Publicada en el Diario Oficial de la Federacion el 26 de Febrero de 1992," .

Field, E. (2007): "Entitled to Work: Urban Property Rights and Labor Supply in Peru*," *Quarterly Journal of Economics*, 122(4), 1561–1602.

Galiani, S., and E. Schargrodsky (2010a): "Property rights for the poor: Effects of land titling," *Journal of Public Economics*, 94(9-10), 700–729.

——— (2010b): "Resource Allocation, Transaction Costs, and Land Property Rights," .

Garfinkel, M. R., and S. Skaperdas (2007): "Chapter 22 Economics of Conflict: An Overview," in *Handbook of Defense Economics*, ed. by S. Todd, and H. Keith, vol. Volume 2, pp. 649–709. Elsevier.

HANSON, G. H. (2005): *Why Does Immigration Divide America? Public Finance and Political Opposition to Open Borders*. Institute for International Economics, Washington, D.C.

HANSON, G. H. (2006): "Illegal Migration from Mexico to the United States," *Journal of Economic Literature*, 44(4), 869–924.

——— (2010): "International Migration and the Developing World," in *Handbook of Development Economics*, ed. by R. Dani, and R. Mark, vol. Volume 5, pp. 4363–4414. Elsevier.

HANSON, G. H., AND C. MCINTOSH (2009): "The Demography of Mexican Migration to the United States," *American Economic Review*, 99(2), 22–27.

HANSON, G. H., AND C. MCINTOSH (2010): "The Great Mexican Emigration," *Review of Economics and Statistics*, 92(4), 798–810.

HANSON, G. H., AND A. SPILIMBERGO (1999): "Illegal Immigration, Border Enforcement, and Relative Wages: Evidence from Apprehensions at the U.S.-Mexico Border," *American Economic Review*, 89(5), 1337–1357.

HEATH, J. R. (1990): "Enhancing the Contribution of Land Reform to Mexican Agricultural Development," .

HECKMAN, J., N. HOHMANN, J. SMITH, AND M. KHOO (2000): "Substitution and Dropout Bias in Social Experiments: A Study of an Influential Social Experiment," *The Quarterly Journal of Economics*, 115(2), 651–694.

HECKMAN, J. J., R. J. LALONDE, AND J. A. SMITH (1999): "Chapter 31 The economics and econometrics of active labor market programs," in *Handbook of Labor Economics*, ed. by C. A. Orley, and C. David, vol. Volume 3, Part A, pp. 1865–2097. Elsevier.

HECKMAN, J. J., AND R. ROBB JR (1985): "Alternative methods for evaluating the impact of interventions: An overview," *Journal of Econometrics*, 30(1âĂŞ2), 239–267.

IBARRARAN, P., AND D. LUBOTSKY (2007): "Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census," .

McKenzie, D., and H. Rapoport (2010): "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks," *Review of Economics and Statistics*, 92(4), 811–821.

Mendola, M. (2008): "Migration and technological change in rural households: Complements or substitutes?," *Journal of Development Economics*, 85(1-2), 150–175.

Mishra, P. (2007): "Emigration and wages in source countries: Evidence from Mexico," *Journal of Development Economics*, 82(1), 180–199.

Morales Jurado, M., and A. Colin Salgado (2006): "Controversias agrarias," *Estudios Agrarios (Procuraduria Agraria)*, 31, 223–234.

Mullan, K., P. Grosjean, and A. Kontoleon (2011): "Land Tenure Arrangements and Rural-Urban Migration in China," *World Development*, 39(1), 123–133.

Munoz-Pina, C., A. De Janvry, and E. Sadoulet (2003): "Recrafting Rights over Common Property Resources in Mexico," *Economic Development & Cultural Change*, 52(1), 129–158.

Munshi, K. (2003): "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market*," *Quarterly Journal of Economics*, 118(2), 549–599.

NACLA (1976): "Harvest of Anger," *Latin America and Empire Report*, 10(July-August).

Norton, E. C., H. Wang, and C. Ai (2004): "Computing Interaction Effects and Standard Errors in Logit and Probit Models," *Stata Journal*, 4, 154–167.

Nuijten, M. (2003): "Family Property and the Limits of Intervention: The Article 27 Reforms ad the PROCEDE Program in Mexico," *Development and Change*, 34(3), 475–497.

Orrenius, P. M., and M. Zavodny (2005): "Self-selection among undocumented immigrants from Mexico," *Journal of Development Economics*, 78(1), 215–240.

Pande, R., and C. Udry (2006): "Institutions and Development: A View from Below," in *Advance in Economics and Econometrics*, ed. by R. Blundell, W. K. Newey, and T. Persson. Cambridge University Press.

Passel, J. S. (2005): "Estimates of the Size and Characteristics of the Undocumented Population," Discussion paper, Pew Hispanic Center, Washington, D.C.

Procuraduaria Agraria (2007): "La organizacion agraria como factor de desarrollo en ejidos y comunidades," *Estudios Agrarios (Procuraduria Agraria)*, 34, 165–178.

——— (2010): "Caracteristicas basicas de los nucleos agrarios certificados," *Estudios Agrarios (Procuraduria Agraria)*, 45, 165–188.

Quesnel, A. (2003): "Poblamiento, regulaciones agrarias y movilidad en el sur del estado de Veracruz," in *Dinamicas de poder y juegos de actores entorno a la tenencia de la tierra*, ed. by A. Q. y. E. V. E. Leonard, pp. 41–71. CIESAS-IRD., Mexico, D.F.

Riveros Fragoso, L. (2005): "La libertad de la transmision en las sucesiones agrarias," *Estudios Agrarios (Procuraduria Agraria)*, 28, 31–48.

Rosenzweig, M. R. (2007): "Education and migration: A global perspective," Yale University.

Sadoulet, E., A. d. Janvry, and B. Davis (2001): "Cash Transfer Programs with Income Multipliers: PROCAMPO in Mexico," *World Development*, 29(6), 1043–1056.

Singh, I., L. Squire, and J. Strauss (1986): "A Survey of Agricultural Household Models: Recent Findings and Policy Implications," *The World Bank Economic Review*, 1(1), 149–179.

Skaperdas, S. (1996): "Contest success functions," *Economic Theory*, 7(2), 283–290.

U.S. Immigration and Naturalization Service (1999): "Statistical Yearbook of the Immigration and Naturalization Service, 1997," Discussion paper, U.S. Government Printing Office: Washington, D.C.

——— (2003): "Estimates of the Unauthorized Immigrant Population Residing in the United States: 1990 to 2000," Discussion paper, Department of Homeland Security.

Velez, F. (ed.) (1995): *Los Desafios que Enfrenta el Campo en Mexico*, Mexico a la Hora del Cambio. Cal y Arena, Mexico.

Winters, P., A. de Janvry, and E. Sadoulet (2001): "Family and Community Networks in Mexico-U.S. Migration," *Journal of Human Resources*, 36(1), 159–184.

Woodruff, C., and R. Zenteno (2007): "Migration networks and microenterprises in Mexico," *Journal of Development Economics*, 82(2), 509–528.

World Bank (1999): "Mexico Ejido Reform: Avenues Five Years Later," Discussion paper, World Bank, Washington, D.C.

——— (2001): "Mexico's Second Agrarian Reform: Implementation and Impact," Discussion paper, World Bank, Washington, D.C.

——— (2010): "World Development Indicators on CD-ROM," .

——— (2011): "Migration and Remittances Factbook 2011," Discussion paper, World Bank, Washington, D.C.

Yang, D. (2008): "International Migration, Remittances and Household Investment: Evidence from Philippine Migrants Exchange Rate Shocks," *The Economic Journal*, 118(528), 591–630.

Yates, P. (1981): *Mexico's Agricultural Dilemma.* Tucson: the University of Arizona Press.

Appendix

7 Theoretical model

7.1 Equilibrium

The decision problem for household member i can be solved by backward induction.

First, consider member i's second-period allocation choice (in case of capture of the land inheritance). Drop the time-subscripts and write off-farm labor supply in terms of in-farm labor supply: $T_{io} = \bar{T} - T_{if}$. Once we do this, the choice variable is only the amount of in-farm labor supply and we can further simplify the notation: $T_{if} = T_i$. Member i faces the following problem:

$$\max_{T_i} \left\{ Y(T_i, L) + w(\bar{T} - T_i) \right\} \text{ s.t. } \begin{cases} \bar{T} - T_i \geq 0 \\ T_i \geq 0 \end{cases}$$

The corresponding first-order conditions are:

$$\begin{cases} Y_1(T_i^*, L) - w + \lambda \leq 0 \ (= 0 \ if \ T_i^* > 0) \\ \lambda \geq 0, \text{ with } \lambda(\bar{T} - T_i^*) = 0 \end{cases}$$

where λ is the Lagrange multiplier associated with the first constrain.

The end-point restriction in assumption 1 ensures that i's in-farm labor supply is strictly positive. However, we could either have an internal solution ($T_i^* = Y_1^{-1}(w)$) or a corner solution ($T_i^* = \bar{T}$). Label i's optimal choice as $T_i^* = \hat{T}$, where $\hat{T} = \min \left[Y_1^{-1}(w), \bar{T} \right]$.

If member i does not capture the land inheritance, then he has access only to migration and so $T_i^* = 0$.

Consider i's first-period decision problem. Again, drop the time subscript and write off-farm labor supply in terms of in-farm labor supply. Member i faces the following maximization problem:

$$\max_{T_i} \frac{1}{N} Y(T_i + \sum_{k\neq i} T_k, L) + w(\bar{T} - T_i) + \delta \left\{ p^i \left[Y\left(\hat{T}, L\right) - w\hat{T} \right] + w\bar{T} \right\}$$
$$s.t. \bar{T} - T_i \geq 0, T_i \geq 0$$

The corresponding first-order conditions are:

$$\begin{cases} \frac{1}{N} Y_1 - w + \delta \left[Y\left(\hat{T}, L\right) - w\hat{T} \right] f'(T_i^*) \left[\sum_{k\neq i} f\left(T_k\right) + f(T_E) \right] \Pi^{-2} p_1^i + \lambda \leq 0 \; (= 0 \text{ if } T_i^* > 0) \\ \lambda \geq 0, \text{ with } \lambda(\bar{T} - T_i^*) = 0 \end{cases} \tag{i}$$

where $Y_1 \equiv Y_1(T_i^* + \sum_{k\neq i} T_k, L)$, $\Pi \equiv f\left(T_i^*\right) + \sum_{k\neq i} f\left(T_k\right) + f(T_E)$, $p_1^i \equiv p_1 \left(\frac{f(T_i^*)}{\Pi}, \theta \right)$, and λ is the Lagrangean multipliers associated with the first and second constraint.

Since the structure of the decision problem is identical for all household members, their optimal choices will also be identical. This, joint to the end-point restriction we made in assumption 1, ensures that i's optimal in-farm labor supply will be strictly positive. Thus, we could either have an internal solution or a corner solution where i devotes the labor supply exclusively to the in-farm activity.

Consider the case of an internal solution. Define the argument of the maximization problem in (i) as W^i, so that the first-order condition for household member i corresponds to equation (i) without the Lagrangian multipliers, which we can recall as

$$T_i^* : W_{T_i}^i(T_i^*, T_{k\neq i}) = 0. \tag{ii}$$

This is the necessary condition for T_i^* to be optimum. The second-order condition is:

$$W_{T_iT_i}^i = \frac{1}{N} Y_{11} + \left\{ \begin{array}{c} p_{11}^i \left[f'(T_i^*)\right]^2 \left[\sum_{k\neq i} f\left(T_k\right) + f(T_E) \right]^2 \Pi^{-2} + \\ + p_1^i \left[f''(T_i) - 2\left[f'(T_i)\right]^2 \Pi^{-1} \right] \left[\sum_{k\neq i} f\left(T_k\right) + f(T_E) \right] \end{array} \right\} a$$

where $Y_{11} \equiv Y_{11}(T_i^* + \sum_{k\neq i} T_k, L)$, $\Pi \equiv f\left(T_i^*\right) + \sum_{k\neq i} f\left(T_k\right)$, $p_{11} \equiv p_{11} \left(\frac{f(T_i^*)}{\Pi}, \theta \right)$, $p_1 \equiv$

$p_1\left(\frac{f(T_i^*)}{\Pi}, \theta\right)$ and $a \equiv \delta\left[Y(\hat{T}) - w\hat{T}\right](\Pi)^{-2}$.

Since Y_{11}, p_{11} and f'' are negative, while p_1 and f' are positive, then $W^i_{T_iT_i} < 0$. So the function W^i is strictly concave and equation (ii) is a sufficient condition for T_i^* to be the maximum.

The pure-strategy Nash equilibrium is the vector of optimal in-farm labor supplies $(T_1^*, .., T_N^*)$ with generic element T_i^* such that equation (ii) is valid simultaneously for all household members. As we noticed above, in equilibrium household members' equilibrium choices will be identical: $T_1^* = T_2^* = .. = T^*$.

7.2 Comparative statics

Notice that the equilibrium condition for household member i is $W^i_{T_i}\left(T_1^*, .., T_N^*; N, L, w, \delta, s, \bar{T}, \theta\right) = 0$. Totally differentiate $W^i_{T_i}$ and assume that $dN = dL = dw = d\delta = ds = d\bar{T} = 0$, while $d\theta \neq 0$. Then the comparative static for household member i is:

$$\frac{dT_i^*}{d\theta} = \frac{\begin{vmatrix} W^1_{T_1T_1} & .. & -W^1_{T_1\theta} & .. & W^1_{T_1T_N} \\ .. & .. & .. & .. & .. \\ W^N_{T_1T_N} & .. & -W^N_{T_N\theta} & .. & W^N_{T_NT_N} \end{vmatrix}}{\begin{vmatrix} W^1_{T_1T_1} & .. & W^1_{T_1T_i} & .. & W^1_{T_1T_N} \\ .. & .. & .. & .. & .. \\ W^N_{T_1T_N} & .. & W^N_{T_NT_i} & .. & W^N_{T_NT_N} \end{vmatrix}} \tag{iii}$$

where all elements are evaluated in correspondence of the equilibrium vector $(T_1^*, .., T_N^*)$

and the generic elements $W^i_{T_iT_i}$, $W^i_{T_iT_j}$ and $W^i_{T_i\theta}$ are:

$$\begin{aligned}
W^i_{T_iT_i} &= \frac{1}{N}Y_{11} + \left\{ \begin{array}{c} p^i_{11}\left[f'(T^*_i)\right]^2 \left[\sum_{k\neq i} f\left(T^*_k\right) + f(T_E)\right]^2 (\Pi^*)^{-2} + \\ +p_1 \left[f''(T^*_i) - 2\left[f'(T^*_i)\right]^2 (\Pi^*)^{-1}\right] \left[\sum_{k\neq i} f\left(T^*_k\right) + f(T_E)\right] \end{array} \right\} a \\
W^i_{T_iT_j} &= \frac{1}{N}Y_{11} + \left\{ \begin{array}{c} -p^i_{11} f'(T^*_i) \left[\sum_{k\neq i} f\left(T^*_k\right) + f(T_E)\right] f'\left(T^*_j\right) f(T^*_i) (\Pi^*)^{-2} + \\ +p_1 \left\{1 - 2\left[\sum_{k\neq i} f\left(T^*_k\right) + f(T_E)\right] (\Pi^*)^{-1}\right\} f''(T^*_j) f'(T^*_i) \end{array} \right\} a \\
W^i_{T_i\theta} &= p_{12} f'(T^*_i) \left[\sum_{k\neq i} f\left(T^*_k\right) + f(T_E)\right] a
\end{aligned}$$

Since in equilibrium $T^*_1 = .. = T^*_N = T^*$, the previous expressions can be simplified significantly: $f\left(T^*_i\right) = f\left(T^*_j\right) = f\left(T^*\right) = f$, $\sum_{k\neq i} f\left(T^*_k\right) + f(T_E) = \Pi^* - f$, $f'\left(T^*\right) \equiv f_T$, $f''\left(T^*\right) \equiv f_{TT}$ $\forall i, j$ and $f(T_E) = f_E$. We also drop the star symbol from Π^*. The previous expressions become:

$$\begin{aligned}
W^i_{T_iT_i} &= \frac{1}{N}Y_{11} + \left\{p_{11}\left(f_T\right)^2 \left(\Pi - f\right)^2 \Pi^{-2} + p_1\left[f_{TT} - 2\left(f_T\right)^2 \Pi^{-1}\right]\left(\Pi - f\right)\right\} a \\
W^i_{T_iT_j} &= \frac{1}{N}Y_{11} + \left\{-p_{11}\left(f_T\right)^2 \left(\Pi - f\right) f \Pi^{-2} + p_1\left[1 - 2\left(\Pi - f\right)\Pi^{-1}\right]\left(f_T\right)^2\right\} a \\
W^i_{T_i\theta} &= p_{12} f_T \left(\Pi - f\right) a
\end{aligned}$$

Consider the denominator in equation (iii). Subtract column (N) from columns (1) to (N-1) and "move out" the common factor a from columns (1) to (N-1). Then add rows (1) to (N-1) to row (N).

Consider the numerator. Extract the common factor from column (i). Then subtract row (i) from all other rows and extract the common factor a from the latter.

$$= -\frac{a^N f_T(\Pi - f)\, p_{12}}{a^{N-1}} \frac{\begin{vmatrix} \psi & 0 & .. & 0 & .. & 0 \\ 0 & \psi & .. & 0 & .. & 0 \\ .. & .. & .. & .. & .. & .. \\ W^i_{T_iT_1} & W^i_{T_iT_2} & .. & 1 & .. & W^i_{T_iT_N} \\ .. & .. & .. & .. & .. & .. \\ 0 & 0 & .. & 0 & .. & \psi \end{vmatrix}}{\begin{vmatrix} \psi & 0 & .. & 0 & W^1_{T_1T_N} \\ 0 & \psi & .. & 0 & W^2_{T_2T_N} \\ .. & .. & .. & .. & .. \\ 0 & 0 & .. & \psi & W^{N-1}_{T_{N-1}T_N} \\ 0 & 0 & .. & 0 & \phi \end{vmatrix}}$$

$$= -\frac{af_T(\Pi - f)\, p_{12}}{\phi} = -\frac{af_T(\Pi - f)\, p_{12}}{Y_{11} + p_1\left[f_{TT}(\Pi - f) - (N-1)(f_T)^2 - 2(f_T)^2 f_E \Pi^{-1}\right] a}$$

where $\psi \equiv p_{11}(f_T)^2(\Pi - f)\Pi^{-1} + p_1\left[f_{TT}(\Pi - f) - (f_T)^2\right]$,

$\phi \equiv Y_{11} + p_1\left[f_{TT}(\Pi - f) - (N-1)(f_T)^2 - 2(f_T)^2 f_E \Pi^{-1}\right] a$ and $W^1_{T_1T_N} = W^2_{T_2T_N} = .. = W^{N-1}_{TN-1T_N}$.

Since p_{12}, Y_{11} and f_{TT} are negative, while f_T and p_1 are positive, then $\frac{dT_i^*}{d\theta} < 0\ \forall i = 1,..,N$.

Since $T_{io} = \bar{T} - T_i$, then $\frac{dT_i^*}{d\theta} < 0$ implies $\frac{dT_{io}^*}{d\theta} > 0$.

Consider the case of a corner solution: all household members devote their entire household labor supply to the in-farm activity ($T^* = \bar{T}$). An increase in land property rights (θ) may not be enough to change the equilibrium choice from corner to internal, so the comparative static will be $\frac{dT_i^*}{d\theta} \leq 0$ and $\frac{dT_{io}^*}{d\theta} \geq 0$.

8 Derivation of the estimator

Re-write the baseline comparison in terms of potential outcomes:

$$\{E[M_i(1,1)|P=1,E=1]-E[M_i(1,0)|P=1,E=0]\}+$$
$$-\{E[M_i(0,1)|P=0,E=1]-E[M_i(0,0)|P=0,E=0]\}.$$

The assumption of random allocation of the program across eligible and non-eligible households lets us manipulate this expression as follows:

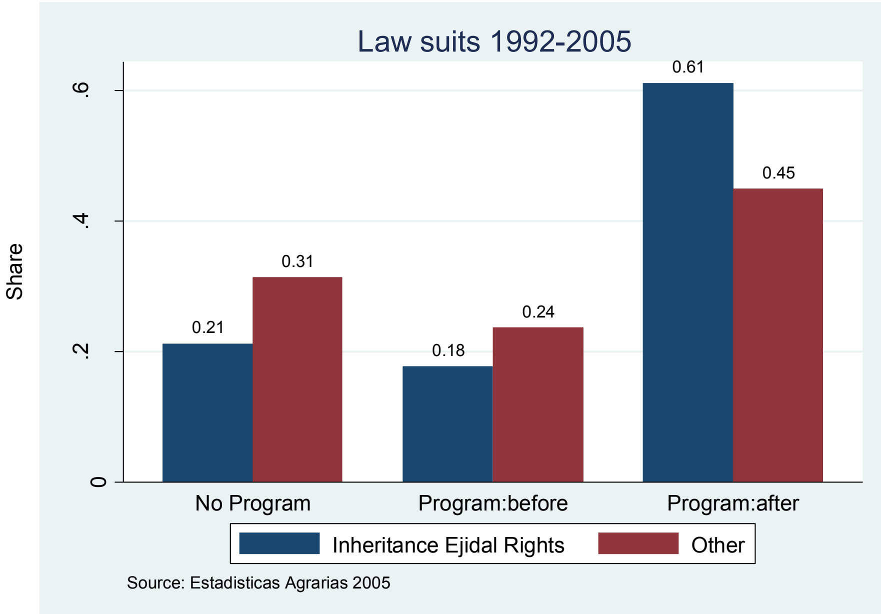
$$\{E[M_i(1,1)|P=1,E=1]-E[M_i(1,0)|P=1,E=0]\}+$$
$$-\{E[M_i(0,1)|P=1,E=1]-E[M_i(0,0)|P=1,E=0]\},$$

which clearly reduces to:

$$\{E[M_i(1,1)-M_i(0,1)|P=1,E=1]-E[M_i(1,0)-M_i(0,0)|P=1,E=0]\}.$$

Figure 1

Law cases concerning land inheritance before and after Procede



Note: the figure shows the differential increase of law suits concerning land inheritance (relative to other categories) after the program took place. See Morales Jurado and Colin Salgado (2006) for details.

Table 1

PANEL A: REASONS TO IMPLEMENT THE PROGRAM		
Sample	Program areas (N=111)	Non-program areas, in process (N=41)
	mean	mean
Tenure security	0.883	0.756
Solve border issues	0.297	0.293
Obey the law	0.153	0.146
Access credit	0.108	0.098
Rent and sell the land	0.108	0.024
Access to Procampo	0.018	0.098
Invest in the land	0.018	0.000
Other	0.000	0.024

PANEL B: REASONS NOT TO IMPLEMENT THE PROGRAM	
Sample	Non program areas, program not even started (N=69)
	mean
Lack of information	0.304
Avoid taxes	0.159
Border issues	0.159
Avoid conflicts between ejidatarios and non-ejidatarios	0.087
They did not summoned us	0.029
Lack of documents	0.043
Avoid land transactions	0.014
No interest in selling and buying land	0.000
Other	0.000

Note: Data from the 1997 community-level ejido survey. Ejidos that had terminated or started to implement Procede were asked the reasons for their decision to implement. Ejidos that had not started to implement the program were asked about the reason for this.

Table 2

DESCRIPTIVE STATISTICS, COMMUNITY-LEVEL

	1994			1997		
	Program mean	No Program mean	Diff	Program mean	No Program mean	Diff
	(1)	(2)	(3)	(4)	(5)	(6)
Log ejido area (ha)	6.85	7.14	*	7.00	7.16	
% urban area wrt ejido area (ha)	3.53	3.41		2.80	2.28	
% parcelled land wrt agr land (ha)	70.84	58.21	***	73.02	59.80	***
Number of ejidatarios [1]	87.01	112.74	**	104.46	108.65	
Number of posesionarios[1]	N/A			9.67	24.87	**
Number of avecindados[1]	73.55	62.91		53.92	45.67	
Ratio avecindados/ejidatario households	0.85	0.67		0.64	0.50	
Average parcelled land per ejidatario (ha)	13.12	11.90		14.69	12.04	
Inequality land[2]	6.03	9.85	*	9.33	10.10	
Common land per ejidatario (ha)	9.84	8.64		9.43	10.56	
Indigenous ejido	0.16	0.11		0.31	0.25	
Membership to ejido union	0.32	0.41		0.25	0.28	
Distance from closest urban centre (km)	N/A			23.93	27.59	
Number of urban centres within a hour	N/A			1.72	1.39	*
At least one irrigation facility	N/A			0.42	0.31	*
At least one storing facility	N/A			0.15	0.19	
Access to paved road	0.35	0.22	**	0.70	0.58	*
% dwellings with electricity	79.79	71.31	*	82.32	80.05	
% dwellings with drinking water	62.21	49.06	**	68.13	54.57	**
% dwellings with drenage	15.19	13.22		14.06	9.41	
Public phone	0.55	0.49		0.61	0.53	
Street lightning	0.69	0.63		0.73	0.72	
Auditorium/assembly hall	0.61	0.44	***	0.64	0.38	***
External boundary problems[3]	0.24	0.59	***	0.12	0.47	***
Internal boundary problem[3]	N/A			0.14	0.18	
Boundary problem in communal land[3]	0.14	0.40	***	0.06	0.09	
Squatting common land[3]	N/A			0.12	0.30	***
Kindergarden[3]	N/A			0.80	0.85	
Primary school[3]	0.96	0.95		0.95	0.96	
Secondary school[3]	N/A			0.44	0.49	
At least one social program	0.57	0.46		0.54	0.54	
At least one environmental problem	N/A			0.42	0.50	
Observations	111	110		111	110	

* significant at 10%; ** significant at 5%; *** significant at 1%. Column (3) reports the significance of the difference (1)-(2). Column (6) reports the significance level of the difference (4)-(5). Definition of "Program" in the text.

[1] *Posesionarios* are households with ejido membership and formal right to land; *avecindados* are households with ejido membership but no formal right to land, although part of them own land illegally; *posesionarios* are households with no ejido membership and no formal right to land, although most of them owns land illegally.

[2] Land inequality measured as the ratio between the biggest and the smallest plot for entitled individuals.

[3] The definition of some variables differ across the two surveys: indigenous ejido (1997: "Are there people who consider themselves indigenous?"; 1994:"Does the majority belong to an ethnic group?"); external boundary problems (1997: "Are there boundary problems with other ejidos or other borderign private properties?"; 1994: "Are there law problems concerning the ejido borders?"); internal boundary problems (1997:"Are there boundary problems between ejidatarios about the division of parcelled land?"; 1994: none); boundary problems related to communal land (1997: "Are there border problems between ejidatarios about the assignment of communal land?"; 1994:"Are there problems concerning the borders of communal land?"); squatting of communal land (1997: "Is there communal land squatted by families without documentation?"; 1994: none); schools (1997:"Does the community have a kindergarden/ primary/secondary school?"; 1994:"Does the community have a school?").

Table 3

PRE-PROGRAM DESCRIPTIVE STATISTICS, HOUSEHOLD LEVEL

	1994 All mean (1)	1997 All mean (2)	1994 Eligible Program mean (3)	1994 Eligible No Program mean (4)	Diff t-stat (5)	1994 Non-Eligible Program mean (6)	1994 Non-Eligible No Program mean (7)	Diff t-stat (8)	Diff-diff t-stat (9)
A: Migration variables									
At least one household member currently living at home has been abroad (last 3 years)	0.04	0.08	0.04	0.04	(0.446)	0.02	0.04	(-1.118)	(1.196)
At least one household head's child is currently abroad	0.12	0.23	0.10	0.15	(-1.439)	0.09	0.12	(-0.527)	(-0.442)
Migrant household (last 3 years)	0.15	0.29	0.14	0.17	(-0.784)	0.11	0.15	(-0.629)	(0.045)
Number of migrants abroad (last 3 years)	0.30	0.72	0.27	0.38	(-1.048)	0.20	0.31	(-0.945)	(0.013)
B: Household composition									
Household head's age	49.85	52.88	51.06	50.83	(0.156)	48.04	47.59	(0.255)	(-0.111)
Household head's sex	0.97	0.97	0.97	0.97	(-0.030)	0.95	0.99	(-1.695) *	(1.397)
Household head's schooling	3.27	3.20	3.34	3.12	(0.784)	3.45	3.24	(0.564)	(0.047)
Average schooling of adult members	4.68	4.66	4.79	4.67	(0.409)	4.48	4.71	(-0.663)	(0.868)
Number of adult members	5.92	6.71	6.14	6.16	(-0.061)	5.31	5.67	(-0.831)	(0.638)
Share females among adult members	0.44	0.37	0.45	0.44	(0.804)	0.44	0.41	(1.432)	(-0.677)
Number of household head's siblings abroad	0.14	0.38	0.11	0.20	(-1.468)	0.09	0.12	(-0.576)	(-0.891)
C: Household assets									
1992 land assets (owned)	11.76	11.76	12.09	12.31	(-0.139)	10.61	11.29	(-0.325)	(0.205)
Hired labor	0.37	0.45	0.42	0.37	(0.950)	0.36	0.28	(1.211)	(-0.356)
Tractor	0.47	0.46	0.56	0.45	(1.750) *	0.49	0.31	(2.092) **	(-0.722)
Pickup	0.32	0.21	0.37	0.30	(1.340)	0.27	0.31	(-0.549)	(1.332)
Machinery	0.59	0.59	0.66	0.55	(1.931) *	0.61	0.50	(1.419)	(0.023)
Cattle	0.47	0.45	0.44	0.54	(-1.835) *	0.39	0.50	(-1.445)	(0.074)
Horses	0.23	0.30	0.25	0.24	(0.380)	0.23	0.20	(0.614)	(-0.266)
D: Land transactions									
At least one land rental transaction (last 2 years)	0.10	0.21	0.12	0.10	(0.877)	0.12	0.07	(1.458)	(-0.434)
At least one plot rented in (last 2 years)	0.06	0.09	0.06	0.06	(0.211)	0.08	0.05	(0.980)	(-0.671)
At least one plot rented out (last 2 years)	0.04	0.09	0.06	0.04	(0.928)	0.04	0.02	(1.093)	(0.052)
Observations	926	926	298	302		169	157		

* significant at 10%; ** significant at 5%; *** significant at 1%. Columns (1) and (2) report sample means from the 1994 and 1997 surveys respectively. Column (5) reports the t-statistic of the difference (3)-(4). Column (8) reports the t-statistic of the difference (6)-(7). Column (9) reports the t-statistic of the difference [(3)-(4)]-[(6)-(7)]. Standard errors associated with the diff-in-mean tests have been clustered at the ejdo-level. Definition of "Program," "Eligible," and household in the text. Land assets measured in National Rainfed Equivalent (NRE) hectares. For a description of the procedure, see de Janvry et al. (1997).

Table 4
HOUSEHOLD MIGRATION, CROSS-SECTION ESTIMATES

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Migrant household			Number migrants	Share migrants
Model:	LPM	LPM	Logit, marg effects	OLS	OLS
	coef/se	coef/se	coef/se	coef/se	coef/se
Program × Eligible	0.115	0.127*	0.119	0.426**	0.075***
	(0.077)	(0.065)	0.067	(0.200)	(0.025)
Program	-0.081	-0.056	-0.074	-0.239	-0.039*
	(0.066)	(0.056)	0.060	(0.183)	(0.022)
Eligible	-0.031	0.104	0.187	-0.055	0.055
	(0.058)	(0.178)	0.192	(0.589)	(0.076)
Household controls					
Land assets		0.000	0.000	0.003	-0.000
		(0.002)	0.002	(0.008)	(0.001)
Household head's age		0.004*	0.006	0.007	0.002*
		(0.003)	0.003	(0.007)	(0.001)
Average literacy adult household members		0.016	0.014	-0.466	-0.036
		(0.119)	0.101	(0.443)	(0.050)
Average schooling adult household members		0.003	0.007	0.020	0.004
		(0.010)	0.011	(0.027)	(0.003)
Share of females among adult household members		-0.197***	-0.226	-0.357*	-0.047*
		(0.060)	0.093	(0.184)	(0.026)
Household size		0.027***	0.026	0.108***	0.003
		(0.009)	0.009	(0.033)	(0.004)
Number of household head's siblings abroad		-0.033	-0.017	-0.108	-0.016
		(0.049)	0.053	(0.143)	(0.016)
Ejido controls					
Log ejido area (ha)		-0.015	-0.014	-0.075	-0.005
		(0.023)	0.024	(0.088)	(0.010)
% common land relative to agricultural land (ha)		-0.001	-0.001	-0.000	-0.000
		(0.001)	0.001	(0.002)	(0.000)
Number of ejidatarios		-0.000	0.000	0.000	0.000
		(0.000)	0.000	(0.001)	(0.000)
Indigenous ejido		-0.159***	-0.188	-0.356***	-0.053***
		(0.041)	0.048	(0.117)	(0.014)
Membership to ejido union		0.022	0.013	0.117	0.006
		(0.045)	0.045	(0.154)	(0.016)
Access to paved road		-0.097**	-0.103	-0.211	-0.036**
		(0.047)	0.045	(0.150)	(0.017)
Constant	yes	yes	yes	yes	yes
Household controls*Eligible		yes	yes	yes	yes
Observations	926	898	898	898	898
Number of ejidos	221	213	213	213	213
Adjusted R-squared	0.001	0.157	0.169	0.174	0.094

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors (in brackets) clustered at the ejido level. Econometric methodology: Linear Probability Model (LPM) or OLS (Column 1-2, 4-5), Logit (Column 3). The marginal effect associated with the interaction term in Column 4 was computed following Norton, Wang and Ai (2004). Definitions of "Migrant household," "Program," "Eligible," and household in the text. Literacy is computed for members currently living at home only.

Table 5

BASELINE ESTIMATES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	Migrant household					Number migrants	Share migrants	Migrant household	Number migrants	Share migrants
Model:	LPM	LPM	LPM	LPM	Logit, marg effects	OLS	OLS	LPM	OLS	OLS
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
PANEL A: HOUSEHOLD MIGRATION, RESTRICTED SAMPLE										
Program × 1997	-0.035	-0.060	-0.062	-0.054	-0.039	-0.175	-0.032			
	(0.062)	(0.060)	(0.069)	(0.062)	0.065	(0.191)	(0.021)			
Timing × 1997								-0.026	-0.071	-0.015
								(0.044)	(0.136)	(0.014)
Observations	452	451	451	451	451	451	451	432	432	432
Number of ejidos	98	98	98	98	98	98	98	93	93	93
Adjusted R-squared	0.042	0.126	0.495	0.156	0.152	0.159	0.141	0.158	0.165	0.148
PANEL B: HOUSEHOLD MIGRATION, PANEL ESTIMATES										
Program × Eligible × 1997	0.112*	0.124**	0.123*	0.121**	0.129	0.348**	0.070***			
	(0.061)	(0.062)	(0.065)	(0.062)	0.056	(0.167)	(0.021)			
Program × 1997	-0.047	-0.053	-0.053	-0.054	0.030	-0.180	-0.037**			
	(0.053)	(0.052)	(0.055)	(0.053)	0.035	(0.143)	(0.018)			
Timing × Eligible × 1997								0.087**	0.242**	0.047***
								(0.041)	(0.103)	(0.014)
Timing × 1997								-0.042	-0.129	-0.025**
								(0.034)	(0.084)	(0.011)
Eligible × 1997	-0.056	-0.055	-0.051	-0.042	0.060	-0.183	-0.042**	-0.042	-0.198*	-0.043**
	(0.043)	(0.042)	(0.044)	(0.041)	0.070	(0.114)	(0.018)	(0.040)	(0.108)	(0.017)
1997	0.166***	0.138***	0.141***	0.135***	0.117	0.466***	0.070***	0.136***	0.477***	0.070***
	(0.041)	(0.040)	(0.041)	(0.038)	0.017	(0.113)	(0.017)	(0.036)	(0.106)	(0.016)
Observations	1 852	1 849	1 849	1 849	1 849	1 849	1 849	1 744	1 744	1 744
	221	221	221	221	221	221	221	209	209	209
Adjusted R-squared	0.030	0.130	0.394	0.113	0.141	0.132	0.093	0.112	0.134	0.094
Constant	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Household controls		yes	yes	yes	yes	yes	yes	yes	yes	yes
Fixed effects			ejido	household		household	household	household	household	household

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors (in brackets) clustered at the ejido level. Econometric model: Linear Probability Model (LPM) or OLS (Column 1-4, 6-10), Logit (Column 5). Details of the various specifications at the bottom of the table are valid for both panels. Marginal effects in Column 5 have been computed following Cornelissen and Sonderhof (2009). Definitions of "Migration household", "Program", "Timing", "Eligible," and household in the text. See Table 4 for the list of household controls.

Table 6

ROBUSTNESS TEST: ANTICIPATORY RESPONSE TO THE PROGRAM

PANEL A: CROSS-SECTION 1994

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:	Migrant household			Number migrants	Share migrants	Migrant household		Number migrants	Share migrants
Model:	LPM	LPM	Logit, marg effects	OLS	OLS	LPM	LPM	OLS	OLS
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
Program × Eligible	0.003	-0.009	0.003	-0.000	-0.001				
	(0.063)	(0.058)	0.068	(0.128)	(0.017)				
Soon-To-Be-Certified × Eligible						0.019	0.012	0.018	0.024
						(0.088)	(0.086)	(0.145)	(0.024)
Constant	yes	yes	yes	yes	yes	yes	yes	yes	yes
Household controls		yes	yes	yes	yes		yes	yes	yes
Ejido-controls		yes	yes	yes	yes		yes	yes	yes
Observations	926	872	872	872	872	866	813	813	813
Number of ejidos	221	210	210	210	210	208	196	196	196
Adjusted R-squared	0.000	0.077	0.142	0.087	0.039	-0.002	0.083	0.084	0.051

PANEL B: PANEL SAMPLE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Migrant household				Number migrants	Share migrants	Migrant household	Number migrants	Share migrants
	LPM	LPM	LPM	Logit, marg	LPM	LPM	LPM	LPM	LPM
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
Program × Eligible × 1997	0.097	0.113*	0.111*	0.119	0.308	0.073***			
	(0.064)	(0.066)	(0.065)	0.060	(0.188)	(0.024)			
Timing × Eligible × 1997							0.084**	0.228**	0.048***
							(0.043)	(0.112)	(0.015)
In-Process × Eligible × 1997	-0.045	-0.036	-0.033	-0.024	-0.125	0.005	-0.021	-0.083	0.006
	(0.094)	(0.090)	(0.089)	0.072	(0.234)	(0.034)	(0.088)	(0.228)	(0.033)
Constant	yes	yes	yes	yes	yes	yes	yes	yes	yes
Household controls		yes	yes	yes	yes	yes	yes	yes	yes
Fixed effects			household		household	household	household	household	household
Observations	1 852	1 849	1 849	1 849	1 849	1 849	1 744	1 744	1 744
Number of ejidos	221	221	221	221	221	221	209	209	209
Adjusted R-squared	0.030	0.130	0.113	0.143	0.133	0.093	0.112	0.134	0.094

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors (in brackets) clustered at the ejido level. Marginal effect associated to Logit in Panel A computed following Norton, Wang, and Ai (2004); while in Panel B we followed Cornelissen and Sonderhof (2009). The definitions of "Migrant household," "Program," "Timing," "Eligible," and household are in the text. All estimates in Panel A include: Program (or Timing, in Columns 7-9), Soon-to-be-certified, Eligible. All estimates in Panel B include: Program*1997 (or Timing*1997, in Columns 7-9), In-process*1997, Eligible*1997, 1997. See Table 4 for the list of household and ejido controls.

Table 7

HOUSEHOLD MIGRATION, IMPACT BY INHERITANCE STATUS

	(1)	(2)	(3)
Dependent variable:	Migrant household		
Model:	LPM		
Sample:	All	Will	No Will
	coef/t	coef/t	coef/t
Program × Eligible × 1997	0.121**	0.039	0.147**
	(0.062)	(0.103)	(0.070)
Program × 1997	-0.054	-0.034	-0.041
	(0.053)	(0.077)	(0.060)
Eligible × 1997	-0.042	0.022	-0.051
	(0.041)	(0.066)	(0.049)
1997	0.135***	0.120**	0.122***
	(0.038)	(0.052)	(0.045)
Constant	yes	yes	yes
Household controls	yes	yes	yes
Fixed effects	household	household	household
Observations	1 849	661	1 178
Number of ejidos	221	149	195
Adjusted R-squared	0.113	0.087	0.132

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors (in brackets) clustered at the ejido level. Sample: all households (Column 1); households with a will (Column 2); households without a will (Column 3). Econometric methodology: Linear Probability Model (LPM). Definitions of "Migrant household," "Program," "Eligible," and household in the text. See Table 4 for the list of household controls.

Table A1

AFTER-PROGRAM DESCRIPTIVE STATISTICS, HOUSEHOLD-LEVEL

	All	Eligible			Non-Eligible			
		Program	No Program	Diff	Program	No Program	Diff	Diff-diff
	mean	mean	mean	t-stat	mean	mean	t-stat	t-stat
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Migration variables								
At least one household member currently living at home has been abroad (last 3 years)	0.08	0.11	0.07	(1.513)	0.07	0.07	(0.031)	(1.044)
At least one household head's child is currently abroad	0.23	0.23	0.25	(-0.212)	0.18	0.27	(-1.479)	(1.178)
Migrant household (last 3 years)	0.29	0.32	0.28	(0.610)	0.23	0.31	(-1.224)	(1.498)
Number of migrants abroad (last 3 years)	0.72	0.77	0.72	(0.272)	0.52	0.83	(-1.511)	(1.538)
B: Household composition								
Household head's age	52.88	53.71	53.89	(-0.136)	51.23	51.15	(0.048)	(-0.128)
Household head's sex	0.97	0.96	0.97	(-0.740)	0.96	0.99	(-1.484)	(0.754)
Household head's schooling	3.20	3.33	3.06	(0.935)	3.47	2.92	(1.507)	(-0.675)
Average schooling of adult members	4.66	4.67	4.68	(-0.067)	4.78	4.48	(0.910)	(-0.881)
Number of adult members	6.71	6.65	6.99	(-1.010)	6.40	6.62	(-0.509)	(-0.221)
Share females among adult members	0.37	0.38	0.38	(0.099)	0.38	0.37	(0.431)	(-0.300)
Number of household head's siblings abroad	0.38	0.38	0.38	(-0.016)	0.28	0.47	(-1.225)	(1.121)
C: Household assets								
1992 land assets (owned)	11.76	12.09	12.31	(-0.139)	10.61	11.29	(-0.325)	(0.205)
Hired labor	0.45	0.44	0.49	(-0.984)	0.43	0.41	(0.270)	(-0.843)
Tractor	0.46	0.56	0.41	(2.511) **	0.47	0.37	(1.141)	(0.621)
Pickup	0.21	0.24	0.17	(1.767) *	0.19	0.25	(-0.878)	(1.724) *
Machinery	0.59	0.69	0.53	(2.647) ***	0.60	0.50	(1.199)	(0.642)
Cattle	0.45	0.40	0.53	(-2.430) **	0.36	0.52	(-2.212) **	(0.417)
Horses	0.30	0.28	0.34	(-1.364)	0.21	0.34	(-2.121) **	(0.921)
D: Land transactions								
At least one land rental transaction (1994-1997)	0.21	0.28	0.19	(2.081) **	0.17	0.17	(-0.008)	(1.407)
At least one plot rented in (1994-1997)	0.09	0.11	0.08	(1.227)	0.07	0.10	(-1.044)	(1.559)
At least one plot rented out (1994-1997)	0.09	0.14	0.08	(1.813) *	0.08	0.03	(1.610)	(0.396)
Observations	926	298	302		169	157		

* significant at 10%; ** significant at 5%; *** significant at 1%. Column (1) reports sample means from the 1997 household survey. Columns (4) reports the t-statistics of the difference (2)-(3). Column (7) reports the t-statistic of the difference (5)-(6). Column (8) reports the t-statistic of the difference [(2)-(3)]-[(5)-(6)]. Standard error associated with the diff-in-mean tests have been clustered at the ejido-level. Definitions of "Migrant household," "Program," "Eligible," and household in the text. All migration indicators (but the number of migrants) are binary variables. Land assets measured in National Rainfed Equivalent (NRE) hectares. For a description of the procedure, see de Janvry et al. (1997). The number of adult members is computed relative to the biological household, i.e., household members currently living at home and children of the household head living outside home.

Table A2

PRE-PROGRAM DESCRIPTIVE STATISTICS, EARLY VS LATE PROGRAM AREAS

	All	Eligible			Non-Eligible			
		Early	Late	Diff	Early	Late	Diff	Diff-diff
	mean	mean	mean	t-stat	mean	mean	t-stat	t-stat
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Migration variables								
At least one household member currently living at home has been abroad (last 3 years)	0.04	0.07	0.02	(2.250) **	0.03	0.01	(0.707)	(1.193)
At least one household head's child is currently abroad	0.09	0.07	0.09	(-0.534)	0.11	0.08	(0.393)	(-0.648)
Migrant household (last 3 years)	0.12	0.14	0.11	(0.542)	0.14	0.10	(0.591)	(-0.169)
Number of migrants abroad (last 3 years)	0.23	0.28	0.21	(0.637)	0.19	0.21	(-0.131)	(0.508)
B: Household composition								
Household head's age	49.36	49.85	51.13	(-0.520)	47.69	47.49	(0.075)	(-0.415)
Household head's sex	0.96	0.97	0.96	(0.606)	0.9	0.99	(-2.097) **	(2.195) **
Household head's schooling	3.49	3.11	3.91	(-1.868) *	3.45	3.59	(-0.255)	(-1.049)
Average schooling of adult members	4.71	4.72	4.91	(-0.447)	4.26	4.80	(-1.159)	(0.611)
Number of adult members	3.30	3.27	3.26	(0.040)	2.90	3.74	(-2.256) **	(1.948) *
Share females among adult members	0.45	0.44	0.47	(-0.949)	0.47	0.43	(1.118)	(-1.568)
Number of household head's siblings abroad	0.10	0.15	0.04	(1.845) *	0.14	0.07	(0.963)	(0.561)
C: Household assets								
1992 land assets (owned)	11.86	11.32	13.66	(-0.973)	8.56	13.18	(-1.848) *	(0.877)
Access to electricity	0.69	0.64	0.75	(-0.991)	0.66	0.71	(-0.451)	(-0.386)
Tractor	0.53	0.59	0.47	(1.191)	0.56	0.46	(0.721)	(0.182)
Pickup	0.34	0.36	0.40	(-0.426)	0.31	0.26	(0.481)	(-0.723)
Machinery	0.64	0.68	0.62	(0.667)	0.67	0.60	(0.642)	(-0.069)
Cattle	0.43	0.40	0.50	(-1.118)	0.38	0.42	(-0.439)	(-0.540)
Horses	0.25	0.25	0.28	(-0.549)	0.32	0.14	(2.558) **	(-2.477) **
Observations	414	142	116		72	84		

* significant at 10%; ** significant at 5%; *** significant at 1%. Column (1) reports sample means from the 1994 household survey. Columns (4) reports the t-statistics of the difference (2)-(3). Column (7) reports the t-statistic of the difference (5)-(6). Column (8) reports the t-statistic of the difference [(2)-(3)]-[(5)-(6)]. Standard errors associated with the diff-in-mean tests have been clustered at the ejido level. Definitions of "Migrant household," "Early," "Late," "Eligible,"and household in the text. Land assets measured in National Rainfed Equivalent (NRE) hectares. For a description of the procedure, see de Janvry et al. (1997).

Table A3

PANEL ESTIMATES, ADDITIONAL CONTROLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	Migrant household					Number migrants	Share migrants	Migrant household	Number migrants	Share migrants
Model:	LPM	LPM	LPM	LPM	Logit, marg effects	OLS	OLS	LPM	OLS	OLS
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
Program × Eligible × 1997	0.112*	0.116*	0.118*	0.117*	0.126	0.352**	0.069***			
	(0.061)	(0.061)	(0.064)	(0.061)	0.054	(0.164)	(0.021)			
Program × 1997	-0.047	-0.040	-0.043	-0.044	0.034	-0.159	-0.033*			
	(0.053)	(0.053)	(0.056)	(0.054)	0.036	(0.140)	(0.018)			
Timing × Eligible × 1997								0.085**	0.249**	0.046***
								(0.041)	(0.104)	(0.014)
Timing × 1997								-0.036	-0.124	-0.023**
								(0.035)	(0.086)	(0.011)
Eligible × 1997	-0.056	-0.060	-0.061	-0.053	0.074	-0.239**	-0.046***	-0.052	-0.252**	-0.046***
	(0.043)	(0.042)	(0.044)	(0.041)	0.072	(0.113)	(0.017)	(0.038)	(0.107)	(0.016)
1997	0.166***	0.137***	0.144***	0.140***	0.120	0.490***	0.071***	0.141***	0.506***	0.071***
	(0.041)	(0.041)	(0.042)	(0.039)	0.018	(0.112)	(0.017)	(0.037)	(0.105)	(0.016)
Constant	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Household controls		yes	yes	yes	yes	yes	yes	yes	yes	yes
Household assets		yes	yes	yes	yes	yes	yes	yes	yes	yes
Fixed effects			ejido	household		household	household	household	household	household
Observations	1 852	1 849	1 849	1 849	1 849	1 849	1 849	1 744	1 744	1 744
Number of ejidos	221	221	221	221	221	221	221	209	209	209
Adjusted R-squared	0.030	0.161	0.399	0.131	0.179	0.162	0.111	0.134	0.166	0.114

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors (in brackets) clustered at the ejido level. Econometric model: Linear Probability Model (LPM) or OLS (Column 1-4, 6-10), Logit (column 5). Marginal effects in Column 6 have been computed following Cornelissen and Sonderhof (2009). Definitions of "Migrant household," "Program," "Timing," "Eligible," and household in the text. See Table 4 for the list of household controls. Household assets are the binary indicators for: use tractor; use machinery; ownership cattle; ownership work animal. All additional controls are included in levels (using their pre-program value) and interacted with the 1997 time indicator.

Table A4

HOUSEHOLD MIGRATION, PANEL ESTIMATES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	Migrant household					Number migrants	Share migrants	Migrant household	Number migrants	Share migrants
Model:	LPM	LPM	LPM	LPM	Logit, marg effects	OLS	OLS	LPM	OLS	OLS
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
PANEL A: ELIGIBLE HOUSEHOLDS										
Program × 1997	0.065	0.077*	0.074*	0.074*	0.079	0.200*	0.037**			
	(0.041)	(0.041)	(0.044)	(0.040)	0.042	(0.118)	(0.015)			
Timing × 1997								0.050**	0.128*	0.023**
								(0.024)	(0.069)	(0.009)
1997	0.109***	0.075***	0.083***	0.085***	0.085	0.246***	0.025***	0.088***	0.251***	0.026***
	(0.026)	(0.025)	(0.027)	(0.025)	0.026	(0.064)	(0.008)	(0.023)	(0.061)	(0.008)
Observations	1 200	1 198	1 198	1 198	1 198	1 198	1 198	1 118	1 118	1 118
Number of ejidos	187	187	187	187	187	187	187	176	176	176
Adjusted R-squared	0.028	0.163	0.389	0.121	0.177	0.149	0.102	0.117	0.145	0.096
PANEL B: NON-ELIGIBLE HOUSEHOLDS										
Program × 1997	-0.047	-0.050	-0.050	-0.049	-0.044	-0.180	-0.035*			
	(0.053)	(0.056)	(0.064)	(0.057)	0.061	(0.151)	(0.018)			
Timing × 1997								-0.038	-0.135	-0.023**
								(0.036)	(0.092)	(0.012)
1997	0.166***	0.141***	0.146***	0.143***	0.123	0.509***	0.071***	0.146***	0.528***	0.072***
	(0.041)	(0.043)	(0.047)	(0.042)	0.038	(0.123)	(0.017)	(0.039)	(0.114)	(0.016)
Observations	652	651	651	651	651	651	651	626	626	626
Number of ejidos	141	141	141	141	141	141	141	135	135	135
Adjusted R-squared	0.033	0.160	0.441	0.189	0.197	0.223	0.197	0.195	0.232	0.207
PANEL C: ALL HOUSEHOLDS										
Program × 1997	0.026	0.037	0.035	0.034	0.042	0.064	0.011			
	(0.035)	(0.036)	(0.038)	(0.035)	0.036	(0.098)	(0.012)			
Timing × 1997								0.022	0.039	0.007
								(0.021)	(0.057)	(0.008)
1997	0.129***	0.096***	0.103***	0.106***	0.098	0.336***	0.041***	0.107***	0.342***	0.041***
	(0.025)	(0.024)	(0.025)	(0.024)	0.024	(0.064)	(0.009)	(0.023)	(0.062)	(0.008)
Observations	1 852	1 849	1 849	1 849	1 849	1 849	1 849	1 744	1 744	1 744
Number of ejidos	221	221	221	221	221	221	221	209	209	209
Adjusted R-squared	0.029	0.160	0.396	0.128	0.171	0.158	0.103	0.127	0.159	0.103
Constant	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Household controls		yes	yes	yes	yes	yes	yes	yes	yes	yes
Fixed effects			ejido	household		household	household	household	household	household

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors (in brackets) clustered at the ejido level. Details of the various specifications at the bottom of the table are valid for both panels. Definitions of "Migrant household," "Program," "Timing," "Eligible," and household in the text. See Table 4 for the list of household controls.

Table A5

EXCLUDE EJIDOS WHICH IMPLEMENTED OR FAILED TO IMPLEMENT THE PROGRAM BECAUSE OF BORDER ISSUES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	Migrant household					Number migrants	Share migrants	Migrant household	Number migrants	Share migrants
Model:	LPM	LPM	LPM	LPM	Logit, marg effects	OLS	OLS	LPM	OLS	OLS
	coef/t	coef/t	coef/t	coef/t	coef/t	coef/t	coef/t	coef/t	coef/t	coef/t
Program × Eligible × 1997	0.134*	0.155*	0.151*	0.141*	0.153	0.489**	0.092***			
	(0.080)	(0.079)	(0.085)	(0.080)	0.073	(0.209)	(0.027)			
Program × 1997	-0.052	-0.062	-0.060	-0.054	0.035	-0.219	-0.044**			
	(0.068)	(0.066)	(0.070)	(0.068)	0.046	(0.168)	(0.022)			
Timing × 1997 × Eligible								0.096*	0.293**	0.059***
								(0.052)	(0.131)	(0.018)
Timing × 1997								-0.036	-0.134	-0.028**
								(0.043)	(0.105)	(0.014)
Eligible × 1997	-0.043	-0.040	-0.034	-0.015	0.096	-0.224	-0.049**	-0.018	-0.223	-0.050**
	(0.053)	(0.052)	(0.054)	(0.051)	0.086	(0.144)	(0.023)	(0.050)	(0.141)	(0.022)
1997	0.172***	0.138***	0.139***	0.129***	0.133	0.516***	0.078***	0.132***	0.520***	0.080***
	(0.050)	(0.049)	(0.050)	(0.046)	0.023	(0.139)	(0.022)	(0.044)	(0.133)	(0.021)
Constant	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Household controls		yes	yes	yes	yes	yes	yes	yes	yes	yes
Fixed effects			ejido	household		household	household	household	household	household
Observations	1 328	1 326	1 326	1 326	1 326	1 326	1 326	1 245	1 245	1 245
Number of ejidos	159	159	159	159	159	159	159	150	150	150
Adjusted R-squared	0.035	0.146	0.402	0.134	0.158	0.151	0.112	0.133	0.152	0.114

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors (in brackets) clustered at the ejido level. Econometric model: Linear Probability Model (LPM) or OLS (Column 1-4, 6-10), Logit (column 5). Sample: exclude ejidos which report having implemented the program because of border issues (46 ejidos) or having failed to implement the program because of border issues and/or disputes between eligible and non-eligible households (16 ejidos). Definitions of "Migrant household," "Program," "Timing," "Eligible," and household in the text. See Table 4 for the list of household controls.

Table A6

PANEL ESTIMATES, NON-AGRICULTURAL LABOR (MEMBERS CURRENTLY AT HOME)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	Non-agricultural status					Number laborers	Share laborers	Non-agricultural status	Number laborers	Share laborers
Model:	LPM	LPM	LPM	LPM	Logit, marg effects	OLS	OLS	LPM	OLS	OLS
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
Program × Eligible 1997	-0.073	-0.035	-0.035	-0.018	-0.011	-0.002	-0.010			
	(0.069)	(0.066)	(0.070)	(0.066)	0.064	(0.141)	(0.026)			
Program × 1997	0.090*	0.063	0.064	0.042	0.039	0.063	0.017			
	(0.055)	(0.051)	(0.055)	(0.051)	0.032	(0.120)	(0.021)			
Timing × Eligible × 1997								-0.006	-0.025	-0.006
								(0.037)	(0.078)	(0.016)
Timing × 1997								0.006	0.032	0.005
								(0.029)	(0.067)	(0.013)
Eligible × 1997	0.075	0.065	0.059	0.078*	0.063	0.105	0.024	0.079*	0.143*	0.027
	(0.048)	(0.045)	(0.049)	(0.046)	0.034	(0.085)	(0.018)	(0.045)	(0.087)	(0.017)
1997	-0.019	-0.026	-0.020	-0.031	0.038	-0.078	-0.008	-0.031	-0.103	-0.010
	(0.038)	(0.035)	(0.038)	(0.034)	0.016	(0.066)	(0.014)	(0.033)	(0.066)	(0.013)
Constant	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Household controls		yes	yes	yes	yes	yes	yes	yes	yes	yes
Fixed effects			ejido	household		household	household	household	household	household
Observations	1 852	1 849	1 849	1 849	1 849	1 849	1 849	1 744	1 744	1 744
Number of ejidos	221	221	221	221	221	221	221	209	209	209
Adjusted R-squared	0.006	0.114	0.187	0.080	0.135	0.070	0.065	0.090	0.075	0.070

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors (in brackets) clustered at the ejido level. Econometric model: Linear Probability Model (LPM) or OLS (Column 1-4, 6-10), Logit (column 5). Definition non-agricultural status: binary indicator taking value 1 if at least one member reports working outside agriculture as primary occupation. Definitions of "Program," "Timing," "Eligible," and household in the text. See Table 4 for the list of household controls.

Table A7

PANEL ESTIMATES, OTHER OUTCOMES

	(1)	(2)	(3)	(4)	(5)	(6)
Model:	LPM	LPM	LPM	LPM	Logit, marg effects	LPM
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
PANEL A: LAND TRANSACTIONS (RENTALS)						
Program × Eligible × 1997	0.037	0.036	0.036	0.042	0.077	
	(0.043)	(0.044)	(0.047)	(0.044)	0.086	
Program × 1997	-0.027	-0.027	-0.026	-0.031	-0.004	
	(0.035)	(0.036)	(0.039)	(0.036)	0.020	
Timing × Eligible × 1997						0.019
						(0.028)
Timing × 1997						-0.005
						(0.024)
Eligible × 1997	-0.025	-0.024	-0.026	-0.036	-0.073	-0.027
	(0.032)	(0.034)	(0.036)	(0.032)	0.056	(0.031)
1997	0.045	0.047	0.048	0.053*	0.030	0.042
	(0.028)	(0.030)	(0.032)	(0.029)	0.011	(0.028)
Observations	1 848	1 845	1 845	1 845	1 845	1 740
Number of ejidos	221	221	221	221	221	209
Adjusted R-squared	0.001	0.002	0.154	0.008	0.022	0.009
PANEL B: WAGE (NON-FAMILY) LABOR						
Program × Eligible × 1997	-0.042	-0.018	-0.034	-0.055	-0.014	
	(0.097)	(0.098)	(0.105)	(0.100)	0.102	
Program × 1997	-0.063	-0.083	-0.070	-0.054	-0.098	
	(0.085)	(0.085)	(0.092)	(0.088)	0.048	
Timing × Eligible × 1997						-0.003
						(0.060)
Timing × 1997						-0.074
						(0.049)
Eligible × 1997	-0.011	-0.018	-0.014	-0.009	-0.033	-0.029
	(0.078)	(0.078)	(0.084)	(0.081)	0.053	(0.077)
1997	0.134*	0.144**	0.140*	0.129*	0.088	0.155**
	(0.070)	(0.070)	(0.076)	(0.074)	0.025	(0.070)
Observations	1 851	1 848	1 848	1 848	1 848	1 743
Number of ejidos	221	221	221	221	221	209
Adjusted R-squared	0.010	0.033	0.221	0.024	0.033	0.030
Constant	yes	yes	yes	yes	yes	yes
Household controls		yes	yes	yes	yes	yes
Fixed effects			ejido	household		household

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors (in brackets) clustered at the ejido level. Econometric model: Linear Probability Model (LPM) or OLS (Column 1-4, 6), Logit (Column 5). Dependent variable: land transactions status (Panel A), wage (non-family) labor status (Panel B). Definition land transactions status: binary indicator taking value 1 if the household rented out or rented in land within the previous 3 years. Definition wage (non-family) labor status: binary indicator taking value 1 if the household hired any non-family member within the previous 24 months. Details of the various specifications at the bottom of the table are valid for both panels. Definitions of "Program," "Timing," "Eligible," and household in the text. See Table 4 for the list of household controls.

Paper 3

Local Elections and Corruption during Democratization: Evidence from Indonesia

Michele Valsecchi*

University of Gothenburg

October 26, 2012

Abstract

In this paper we ask whether the direct election of the local government increases accountability and decreases corruption. In order to identify the causal effect of direct elections, we exploit the gradual introduction of local elections in Indonesia and a novel dataset of corruption events that covers all districts during the period 1998-2008. We find that direct elections increase the number of corruption crimes by about half the pre-election average. We also find that embezzlement practices dominate all other types of corruption activities.

Key words: Political Decentralization; Political Institutions; Elections; Corruption.

JEL Classification codes: D02, D72, D73, H83, K4, 017, P16.

*I am grateful to Rocco Macchiavello and Rohini Pande for many useful comments. Special thanks to Ola Olsson and Måns Söderbom for many useful discussions. The usual disclaimers apply. Email: michele.valsecchi@economics.gu.se

1 Introduction

In this paper we ask whether a specific political institution, namely the appointment of the executive through direct elections, causes less or more corruption than the appointment through indirect elections.

In order to identify the causal effect of direct elections, we exploit the gradual introduction of a legislative reform across Indonesian districts. The legislative change provides the local electorate with the power to elect the district head directly rather than through representation by the local parliament. This institutional change is salient because the district head is responsible for the provision of local public goods. Elections are widely regarded as a powerful disciplining and selection device and therefore constitute a corruption-reducing mechanism which is well worth evaluating. Indeed, the reform was introduced primarily because many observers had reported widespread vote-buying practices between district heads and district parliaments.

Our measure of corruption is based on novel data on corruption prosecutions in Indonesia. Our dataset provides several advantages relative to the typical corruption measures used in the literature. First, corruption prosecutions constitute "hard" evidence of corruption, which makes them more reliable than, e.g., experts' surveys or perceived corruption measures. Second, they cover the entire universe of Indonesian districts, and therefore provide better coverage than typical measures generated by sectorial studies or randomized interventions. This is especially attractive because Indonesia is the fourth most populous country in the world. Third, they provide an encompassing view of corruption activities: there is no possibility of mis-measurement due to, say, officials switching from one type of corruption to another.[1] Fourth, they provide a long time span: more than ten years of data in a newly

[1] One advantage of this measure of corruption is that it is likely to go beyond, e.g., petty corruption. Over-focus on petty corruption may be instead the problem with other corruption measures like cross-checking (Banerjee, Mullainathan, and Hanna (2012):47). It is also less likely to be biased by media reports, as households' perception measures can be expected to be to a large degree. In addition, it is likely to be more responsive to changes in actual corruption than, again, households' perceptions: Olken (2009) finds that, although the correlation between actual and perceived corruption is positive, "increasing the missing expenditure measure by 10 percent is associated with just a 0.8 percent increase in the probability a villager believes that there is any corruption in the project" (p.951). Finally, differently from perception-based measures, our corruption measure also includes details about the type of corruption observed and so provides a chance to evaluate its

democratized developing country implies room for the study of several institutional features.

Using the typology in Persson and Tabellini (2004), the introduction of the direct election of the district government constitutes a change in the form of government, from a parlamentarian to a presidential system. A commonly held assumption (Persson and Tabellini (2004) and references therein) is that local deputies have better information on the government than citizens do. If local deputies were perfectly accountable to the citizens or had similar preferences, then a shift to a presidential form of government would unambiguously decrease its accountability to the citizens. However, local deputies are unlikely to be held perfectly accountable to the citizens. In addition, they may be reasonably thought to value eventual private gains from public office as much as government members may do. Thus, collusion between deputies and government members under the parlamentarian system may imply that a shift to a presidential form of government increases its accountability to the citizens (Persson, Roland, and Tabellini (1997)).

Our paper contributes to the empirical literature on political institutions and economic outcomes. Cross-country evidence on the relationship between form of government and corruption has been inconclusive. Persson, Tabellini, and Trebbi (2003) find some evidence that presidential systems are associated with less corruption, but the relationship holds only for "good democracies". In contrast, Kunicova and Rose-Ackerman (2001) find that presidential systems are associated with more corruption. During recent years researchers have tried to complement cross-country evidence with within-country studies whenever specific contexts provided convincing identification strategies. So far, attention has been mainly paid to term limits (Besley and Case (1995), Dal Bot' and Rossi (2011), and Ferraz and Finan (2011)) and information (Besley and Burgess (2002), Ferraz and Finan (2008)). The only papers some close to ours concern the introduction of village elections in China (Zhang, Fan, Zhang, and Huang (2004), Gan, Xu, and Yao (2007), Luo, Zhang, Huang, and Rozelle (2007), Shen and Yao (2008), Martinez-Bravo, Miquel, Qian, and Yao (2012)). The main differences between the two contexts concern: the country-level political system (non-democracy in China, young democracy in Indonesia); the level of election (village in China, district in Indonesia, where

efficiency implications and discuss policies aimed at reducing it.

districts typically include about 100 villages); the pre-election selection mechanism (appointment from upper-government tiers in China, elections from local parliament in Indonesia). While the difference in the nature of the country-level political system matters primarily for the external validity of the results, the differences in terms of administrative level of the elections and in terms of origin of the shift to direct elections matters directly for the interpretation of the results and the relationship with the rest of the literature. First of all, since villages are very small units, one can interpret the effects of Chinese local elections as the outcome of an increase in village-level monitoring and therefore relate to other studies like, for example, Reinikka and Svensson (2004) and Björkman and Svensson (2009). On the contrary, the relationship linking elections to citizens' monitoring is not as straightfoward in case of elections covering 100 or more villages. Second of all, a shift from appointment from upper government tiers may be very different from a shift from appointment from local parliament, particularly in terms of leaders' selection.

Our paper also contributes to the recent micro-literature on corruption (see Svensson (2005), Banerjee, Mullainathan, and Hanna (2012), Olken and Pande (2012) and Zitzewitz (Forthcoming) for some excellent surveys) by providing evidence on one of its institutional determinants and by documenting the prevalence of embezzlement over other types of corruption, such as, among others, bribes. Surprisingly, we find that the introduction of direct elections increases corruption, rather than decreasing it.

2 Context and theoretical framework

The administrative structure in Indonesia is composed of three layers: the central government, the provinces, and the districts. The district administration is responsible for the provision of all local public goods. It is divided into a district parliament (DPRD) and a district government (composed of a district head and a vice). The district parliament is elected (since 1999). Until 2004 the district parliament appointed the district head. Media and many policy-makers observed that the power to appoint and dismiss the district head provided to the local parliaments had favored collusion between the two and had led to widespread vote-buying and corruption. Therefore, in 2004 the central government passed a law shifting the power to elect the district head from the local parliaments to the local electorates. In the rest of the paper we will refer to the appointment of district heads by citizens through the election of local parliaments as *indirect elections*, whereas the appointment of district heads by citizens without intermediation will be referred to as *direct elections*.

There are two features of the Indonesian context that are relevant for our purposes. First, under indirect elections the local government should be accountable to the citizens through political representation in the local parliament, i.e., the local government is accountable to the local parliament and the local parliament is accountable to the citizens. However, the elections for the local parliament are over-shadowed by the national elections since the two take place at the same time. Thus, the local government may effectively be accountable only to the local parliament under indirect elections, and to the citizens under direct elections. Second, one of the main differences between the central and the local governments is that the latter have almost no authority to set tax rates. The average share of district revenues arising from own sources, like taxes on economic activities, is only about 15 percent.[2] The rest of the revenue comes from transfers from the central government. The local parliament can perfectly observe these transfers since it must approve the annual budget. In this context it seems reasonable to assume that the local parliament has as informational advantage over the citizens.[3]

If local deputies had the same preferences as the citizens, then under direct elections we would expect the district head to exploit the asymmetric information vis-a-vis the citizens and divert more resources for private purposes than she would have under indirect elections (Besley (2005), Besley and Smart (2007), Gadenne (2010)). However, once we allow local deputies' preferences to differ from the citizens' and, in addition, we allow them to collude with the district head, things become much more complicated. Apart from the theoretical political economy literature mentioned in the previous section, there are some relevant contributions stemming from the mechanism design literature. Baliga and Sjöström (1998) suggest a moral

[2] Author's tabulations based on the 1995-2006 budget data from the Ministry of Finance.

[3] Gadenne (2010) provides evidence from Brazil that whenever the revenues of local governments are primarily given by transfers rather than taxes, the local government performs strictly worse.

hazard mechanism in the Industrial Organization literature that, applied in this context, would suggest that the district head would divert more resources under direct elections than under indirect elections. In contrast, however, Mookherjee and Tsumagari (2004) focus on an adverse selection mechanism and conclude that direct elections would unambiguously lead to less divertion than indirect elections.[4]

3 Construction of the corruption database

Our measure of corruption is based on documents on corruption cases prosecuted or coordinated by the General Attorney Office (AGO).

The AGO is "responsible for investigating certain types of cases, bringing prosecutions, playing an intermediary position between the investigation process and the trial process, and ensuring the enforcement of judicial orders and decisions of final and conclusive effect. (..) it is the institution that determines whether a case should proceed to Court based on admissible evidence" (General Attorney Office (2011):7).

Following a recent improvement in transparency, the AGO has made available a description of virtually all corruption cases in Indonesia. The documentation includes a description of the case, a description of the accusation, the district attorney office prosecuting the practice, the stage of the prosecution, and several demographic characteristics of the person accused. In order to operationalize the information included in this documentation we extract location of the corruption event, date (or time frame) of the corruption event, whether the case concerns primarily the public or private sphere, what sector the case concerns about, and what type of criminal case.

Table 1 includes some descriptive statistics on the corruption cases that we coded. Out of an initial sample of 1,365 corruption cases, we drop 33 cases due to missing or unclear location, 114 cases due to being located in provinces with special status,[5] 247 cases due to

[4] Again, Mookherjee and Tsumagari (2004) discuss their theoretical framework relative to the IO literature. Hence, this is our interpretation of their result in this context. See Mookherjee (2006) for an excellent survey on decentralization and collusion within the mechanism design literature.

[5] They are Jakarta, Aceh, Papua, and Papua Barat (previously called Irian Jaya Barat).

missing time references, and 2 cases due to them dating back to years preceding 1998, which is when the first significant anti-corruption legislation became law.

The final sample consists of 985 cases for which we have (at least) location and time references. Among them, 212 (or 21 percent) span more than one year. In order to keep the descriptive statistics consistent with the econometric analysis, we duplicate these cases for each year in which they occured. The final dataset includes 1,396 corruption events: 133 are classified as concerning the private sphere, and 1,289 are classified as concerning (at least partially) the public sphere.[6]

We further decompose cases by type of corruption. Cases of embezzlement refer to instances where the suspect misuses or appropriates part or all the funds that the local government has placed in their care. The typical conflict of interest case refers to instances where the suspect allegedly sets up an auction that benefits some specific parties. Among the cases that recur frequently within the private sphere, we have hazard, which typically refers to fishermen using illegal devices (explosive) to fish, and illegal practice, which typically refers to lack of documents to carry out a private business or smuggling goods across the border.

The most common type of criminal case is embezzlement (75.6%), followed by conflict of interest (12.0%), fraud, distribution, extortion, bribery, illegal practice and hazard. Embezzlement is possibly even more dominant among cases concerning the public sphere (80.8%). Among cases concerning the private sphere it is the second most common type (26.3%) together with hazard, after illegal practice (38.3%).

The dominant role of embezzlement relative to other corruption activities is very interesting considering that the literature on corruption has largely focused on bribes notwithstanding the broad definition of corruption as "abuse of public office for private gain" (Olken and Pande (2012), Banerjee, Mullainathan, and Hanna (2012)). Our data suggest that the focus on bribes has come at the expense of embezzlement.[7] Overlooking embezzlement practices may be dangerous, because the relative economic theory (and therefore the policy implication) is likely to be different. For example, since bribes typically involve a private counterpart to the public of-

[6] With respect to the original 1,006 cases, 115 are classified as concerning the private sphere, and 891 are classified as concerning (at least partially) the public sphere

[7] See Reinikka and Svensson (2004), Reinikka and Svensson (2011) and Olken (2009) for two exceptions.

ficial, bribe-reducing schemes may provide incentives to private individuals to report requests for bribes. In contrast, embezzlement may not involve any private counterpart and therefore may require other strategies to tackle the issue.

In the rest of the analysis we focus on public cases for three reasons. First, corruption practices may be structurally different across the private and public sphere, which implies that an aggregate analysis may be misleading. Second, cases concerning the public sphere seem more appropriate for the study of the impact of direct elections on corruption. Third, cases concerning the public sphere constitute a large majority of the recorded cases.

Panel A disaggregates the public sphere cases by sector: the most common sector is administration (54.4%), followed by education (11.7%), food distribution and health care (5.5%), infrastructures (5.5%), and agriculture (4.0%).[8] The corruption-based ranking of the sectors is similar when we restrict our attention to embezzlement.

Panel B, Columns 1-6, shows the distribution of the corruption cases over time. We record very few cases during the years following the inauguration of the anti-corruption strategy (Law 16/1999, then modified in 2002). The proportion of (public sphere) cases per year increases progressively reaching 8.8% in 2005, 11.4% in 2006, 17.8% in 2007, and 20.4% in 2008, and decreasing to 13.4% in 2009, 4.4% in 2010, and 0.2% in 2011. The late decrease in corruption cases is most likely driven by the data collection process: since it typically takes typically between a few months and 2-3 years years to detect a corruption event,[9] it is not strange to observe relatively few cases for recent periods since we coded our sample in the autumn 2011. In the rest of the paper we will restrict our analysis to corruption events that took place during the period 1998-2008.

The next step is to generate a measure of corruption at the district-year level. Given the abundance of districts in Indonesia and the relatively long time span under investigation, we collapse the data at the district-year level and consider a simple binary variable indicating whether a district experienced one or more corruption events in a specific year. The second

[8] For 10.6% of the cases we lack enough information to pin down the exact sector. We feel that this lack of information is not serious enough to drop the observations though.

[9] The mean number of years to detect a crime is 2, while the median is 2.6 and the standard deviation is 2.3.

set of descriptive statistics in Panel B shows how the average number of corruption crimes per district evolves over time. Corruption cases increased steadily over the decade, yet declined in 2008, 2010 and 2011, again, presumably due to the data collection process.

The last two columns show the pattern of prosecutions, i.e., the number of prosecutions per district in a given year and the number of districts with at least one prosecution. The pattern suggest that the first prosecutions started much later than the first recorded corruption crimes. In particular, the timing of the first prosecutions is contemporaneous to the constitution of the Anti-Corruption Commission in 2004, which national and international observers praised for having boosted anti-corruption activities in Indonesia.

4 Identification strategy

Law 32/2004 modifies the selection mechanism of district heads (and vices) by requiring that citizens vote for them. The legislative change was implemented gradually across districts. The reason for this was that elections were held only once the mandate of a ruling district head expired, and the expiration date varied across districts due to, e.g., year of formation of the districts (districts formed quite continuously since 1956), natural deaths of the district heads, and district heads running for governor or for the national parliament. In our sample, 178 districts held elections in 2005, 59 in 2006, 32 in 2007, 113 in 2008, and 32 in 2009 or later. The timing of direct elections in Indonesia has already been used as a source of exogenous variation by Skoufias, Narayan, Dasgupta, and Kaiser (2011)[10] and Burgess, Hansen, Olken, Potapov, and Sieber (2012).[11] Nonetheless, we run an informal test of the quasi-random assignment of the timing of gradual elections across districts by trying to predict a wide range of village characteristics *before* the direct elections were introduced. The econometric specification corresponding to this test is the following:

$$c_{ki} = \mu + \theta D_{ki} + v_{ki},$$

[10] Skoufias et al. (2011) show that the only determinant of the timing of direct elections that is statistically significant is the expiration of the mandate of the previous district head.

[11] See also Fukumoto and Horiuchi (2011) for a similar research design using Japanese municipalities.

where c_{ki} is the characteristic of village k in district i, D_{ki} is the year of the direct elections (2005,..,2009), and v_{ki} indicates the error term clustered at the district level. If the timing of direct elections was uncorrelated with district characteristics, then θ would be small and statistically insignificant. Data on village characteristics come from a village census collected just before the first direct elections.[12] Table 2 shows the coefficient estimates. The timing of direct elections cannot predict any of the district characteristics except, marginally, the average village area and the per capita oil and gas transfer. Otherwise the coefficient estimate is always small and statistically insignificant.

In this paper we take advantage of this heterogeneity in timing to identify the impact of direct elections in two distinct ways. The first identification strategy considers all districts and makes use of a rather standard Difference-in-Difference (DD) strategy with many periods. Our outcome of interest is local corruption. Throughout the paper we use corruption prosecutions as a measure of corruption. Other works using the same idea are Fisman and Gatti (2002b) and Glaeser and Saks (2006), who study the determinants of corruption across US states, and Fisman and Gatti (2002a), who investigate the determinants of corruption across countries. Differently from them, we have information not only on the date of prosecution of each corruption event, but also on the date of the crime itself. Our dependent variable is the number of corruption crimes *committed* in a given year, rather than the number of corruption crimes *prosecuted* in a given year. This makes our measure not only more accurate, but, as we will see shortly, also opens up ways to improve our identification strategy.

The baseline econometric specification is the following:

$$y_{ijt} = \alpha_{1i} + \beta_1 D_{ijt} + \gamma_1 E_{it-1} + \Phi_{jt} + \varepsilon_{1ijt}, \tag{1}$$

where y_{ijt} is the number of corruption events committed in district i at time t, D_{ijt} is a binary variable indicating whether the district elections have taken place, E_{it-1} is a binary variable indicating the year before the elections, α_{1i} indicates the district fixed effects, Φ_{jt}

[12] The village census is the PODES 2006, which was collected by the Indonesian National Institute of Statistics in May-June 2005.

is a vector of region-year fixed effects, and ε_{1ijt} is the error term clustered at the district level[13]. The coefficient associated with the pre-election year (γ_1) captures the possible impact of pre-election campaigning on the number of corruption events. The coefficient associated with direct elections (β_1) captures the impact of direct elections on the number of corruption events as long as there are no omitted variables that vary over time (within regions) and are correlated both with the timing of direct elections and with local corruption.

The second identification strategy exploits some additional features of the legislative change. Following the approval of Law 32/2004 (December 2004), the central government postponed all elections in late 2004 and early 2005 to June 2005 to allow time to prepare the elections.

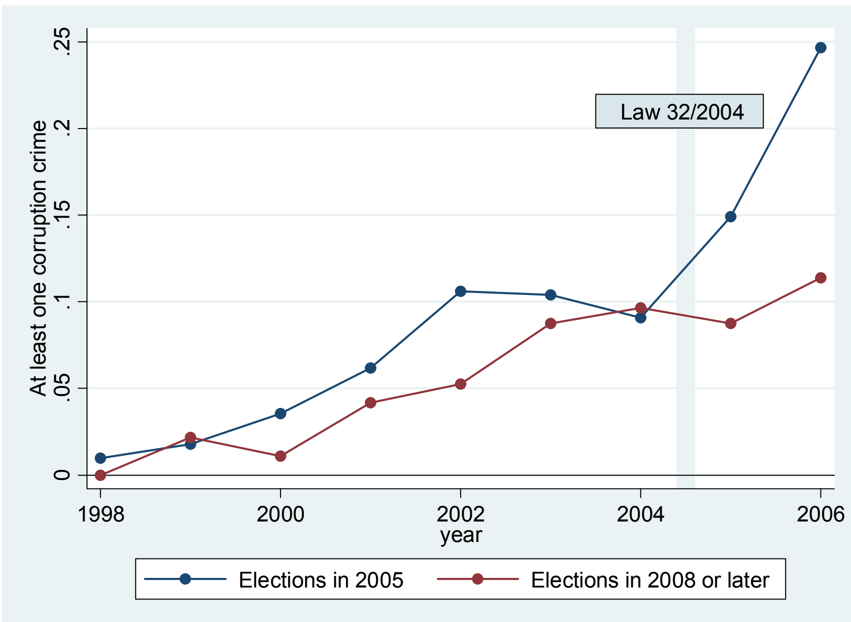


Figure 1: Corruption over time in districts with early and late elections

There are three interesting features of this legislative change: it was discussed and approved in the national parliament in a relatively short time; it was approved in the last

[13]Since many districts split during the period under investigation we cluster the error term according to the district borders as they were in 1999, well before the direct elections were introduced.

quarter of 2004, i.e., after the approval of the 2005 district budget and expenditure plans; and it required incumbents aiming to run for re-election to hand over their seats at least six months before elections to a caretaker appointed by the Ministry of Home Affairs. These three features imply that district governments facing elections in 2005 had very limited opportunities to "anticipate" them (i.e., to modify the district expenditure in order to get re-elected) compared to those facing elections in 2006 or later.

We take advantage of this aspect by comparing districts facing elections in 2005 (treatment group) to those facing elections in 2008 or later (control group). This restriction is also convenient as it allows us to visualize the evolution of corruption in treatment and control districts. Figure 1 suggests that the corruption levels of the two groups are very similar up until the introduction of the election, which is associated with a stark increase in corruption levels in the treatment group. The econometric specification associated with this identification strategy is the same as specification (1) except for the exclusion of the pre-election dummy.

5 Results

5.1 Baseline results

Table 3 shows the coefficient estimates associated with the "restricted sample."[14] Direct elections are associated with a significant increase of about 9 percentage points in the likelihood of having at least one corruption event, i.e., about 90 percent of the pre-election average (Panel A, Columns 1-5). The magnitude of the increase is consistent with the finding that direct elections increase the number of corruption events by about 0.200, i.e., about 100 percent of the pre-election average. Decomposing corruption cases by type yields the following results: direct elections are associated with more embezzlement cases (Panel B), more cases of conflict of interest (Panel C), and more cases of conflict of interest, bribery, and extortion (Panel D). Next, we estimate the impact of direct elections on our treatment group before (1999-2004),

[14] Throughout this section we will discuss primarily the coefficient estimates associated with the specification with region-year fixed effects (Columns 4 and 9 in most of the tables). The results are typically very similar to those with year fixed effects and to those with province-year fixed effects, although the latter are typically less precisely estimated due to the relatively small amount of observations identifying the coefficient of interest.

during (2005) and after elections (2006). The coefficient estimates (Figure 2) substantially mirror the descriptive statistics (Figure 1) and suggest that the bulk of the impact took place the year after elections.

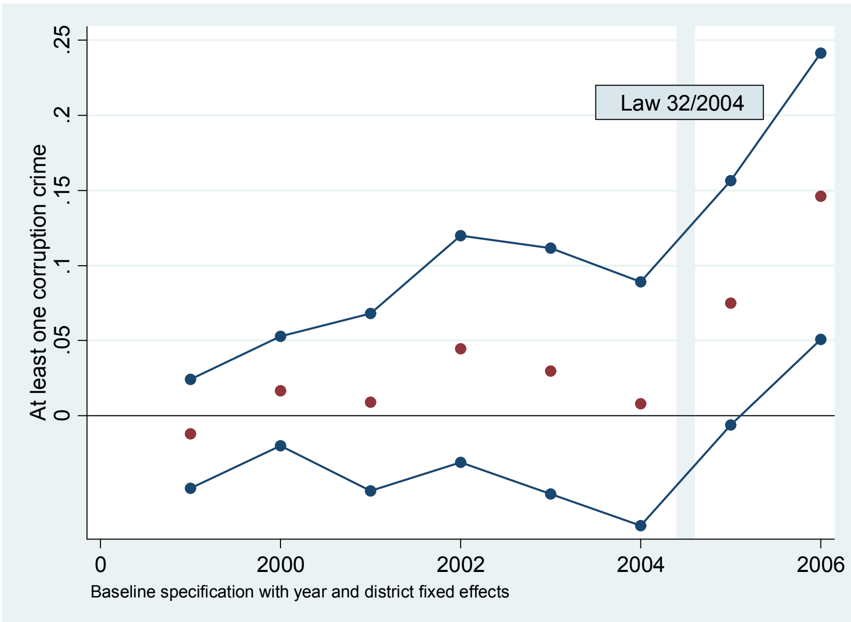


Figure 2: Corruption over time (coefficient estimates)

Table 4 shows the coefficient estimates associated with specification (1). In contrast to the previous estimates we include all districts and control for anticipatory behavior (i.e., campaigning or last-term budget appropriation) by including a binary indicator for the year preceding the elections. The impact of direct elections is again positive, although the magnitude is smaller (about 5 percent, i.e., half the pre-election average) and the coefficient is not always precisely estimated. This holds true even when we consider the number of corruption crimes (Columns 6-10), embezzlement cases (Panel B), or other cases (Panel C and D). The coefficient estimate associated with the pre-election dummy (not reported) is generally positive, insignificant and half the coefficient of interest.

Table 5 shows the coefficient estimates associated with the decomposition of the effect into election year and following years. The coefficient estimates associated with all corruption crimes (Panel A) and embezzlement cases (Panel B) suggest that the magnitude of the impact is rather consistent over time, although the coefficients tend to be precisely estimated only for election years. On the other hand, the coefficient estimates associated with less frequent corruption types (Panel C and D) are positive for elections years and close to zero afterwards. This may explain the lack of significance of the second set of coefficient estimates in Panel A.

In order to test whether the impact of direct elections is driven by some particular area of Indonesia, we re-estimate specification (1) dropping one region/island at a time (Jawa, Kalimantan, Nusa Tenggara and Maluku, Sulawesi, and Sumatera). Table 6 shows the coefficient estimates associated with this robustness check. The impact of direct elections is positive and relatively large across all sample restrictions (it ranges from 0.036 excluding Nusa Tenggara and Maluku to 0.079 excluding Sulawesi).

5.2 Increase in corruption or increase in law enforcement?

The main challenge with the use of corruption prosecutions data is that they (may) capture not only differences in corruption but also differences in law enforcement across districts. In our context this constitutes a measurement error in the dependent variable. This is, however, innocuous for our identification strategy as long as direct elections have no impact on law enforcement at the district level. This could be the case if, for example, direct elections provided a voice to the district electorate, which in turn managed to influence district prosecutors through local newspapers. In a recent work on the US judicial system, Lim, Snyder and Strömberg (2012) find evidence supporting the view that local media can influence local judges. However, they also find that the relationship is driven by elected judges and does not hold for appointed judges, which is exactly how district prosecutors are selected in Indonesia.[15] Corruption prosecutions are typically initiated by district prosecutors, who depend on the provincial prosecutors rather than on the district governments or the district electorate; and provincial prosecutors depend on the General Attorney rather than on the

[15] See also Gordon (2009) for a recent study of prosecutions in the US.

province government or the province electorate.[16]

However, one may still suspect that informal or illegal connections between the district prosecutors and the district government changed with the introduction of direct elections and therefore led to a (positive or negative) change in enforcement, or that direct elections increased the pressure of media and civil society, which in turn may have affected the activity of the prosecutors. We will henceforth refer to this possibility as the "enforcement" channel. In order to rule out the enforcement channel we take the following steps.

The first step is to estimate the impact of direct elections together with a factor that unambiguously increases corruption but not law enforcement.[17] One obvious candidate is a revenue windfall: an increase in government revenue should increase the opportunities for (and the return from) imbezzlement,[18] while having no obvious impact on corruption enforcement.[19] In a companion paper Olsson and Valsecchi (2012), we study the impact of the redistribution of oil and gas revenues across Indonesian districts on a wide range of public goods. In the present paper we exploit the oil and gas revenue transfers to test whether the impact of direct elections on corruption prosecutions captures an increase in corruption or an increase in enforcement. In particular, we estimate the following econometric specification:

$$y_{ijt} = \alpha_{2i} + \beta_{21} D_{ijt} + \beta_{22}(D_{ijt} * windfall_{ij}) + \gamma_2 E_{it-1} + \Phi_{jt} + \varepsilon_{2ijt}, \qquad (2)$$

where $windfall_{ij}$ is the amount of per capita transfers from oil and gas revenues for district i in region j. The coefficient associated with the interaction term (β_{22}) captures the impact of the resource windfall together with the introduction of direct elections. A positive coefficient

[16] Corruption prosecution in Indonesia works as follows: the General Attorney and his staff coordinates the provincial offices, which in turn coordinate the district offices; prosecutions start from investigations by the police or direct complaints from the citizens; on the basis of this preliminary evidence prosecutors decide whether a case is worth further investigation; once they gather enough evidence they send a letter of indictment to the district court office; cases are decided at the district offices; once the verdict has been reached, either the prosecutor or the defendant can bring the case to a higher level (appeal, cassation) in provincial or central courts.

[17] We wish to thank Rohini Pande for this suggestion.

[18] In principle, a greater district government revenue could decrease corruption by increasing officers' salaries. However, in Indonesia officers' salaries are determined and paid by the central government.

[19] In principle, a greater district government revenue could affect enforcement through greater resources allocated to the district attorneys. However, in Indonesia district attorneys are paid by the higher tiers of the AGO structure.

estimate associated with the interaction term should reassure us that the enforcement channel is not driving our results. Table 7 shows the results associated with this robustness test. Since the data on resource-related transfers to district governments date back to 2001, i.e., before some districts formed or split, the sample associated with this analysis is smaller than the one used previously. Panel A shows that this sample restriction reduces the precision (but not the magnitude) of our baseline estimates. Panel B shows the coefficient estimates associated with direct elections and resource abundance: the coefficient of interest is positive and highly significant in all specifications. The magnitude ranges from 0.248 to 0.288 (Columns 1-5) and from 3.515 to 3.832 (Columns 6-10). This implies that an increase in resource transfers of one standard deviation, which equals 0.31 or 310,000 IDR (Table 2), together with direct elections, increases the likelihood of having at least one corruption crime by 76-89 percent (i.e., about 76-89 percent of the pre-election mean) and the number of corruption crimes by 109-119 percent (i.e., about 50-60 percent of the pre-election mean).

The second step we take to rule out the enforcement channel is to estimate the impact of direct elections on the number of corruption crimes *prosecuted* (rather than committed) at time t. In Section 2 we observed that it takes some time to detect a corruption event (the median number of years is 2, while the mean is 2.6), i.e., the crimes prosecuted at time t typically concern events that happened at time t-2. Hence, if direct elections increased mainly law enforcement, the impact on corruption crimes *prosecuted* should be strictly greater than the impact on corruption cases *committed* at a given point in time. If, on the other hand, direct elections truly increased corruption, then the impact on corruption cases prosecuted at a given point in time should be strictly lower than the impact on cases committed. Table 8 shows the coefficient estimates associated with this falsification experiment. The coefficient estimate associated with direct elections is very small and is never significant across all specifications. We interpret such evidence as highly supportive of the main message of the paper.

As an alternative robustness check to the previous falsification experiment we also re-estimated specification (1) controlling for the number of corruption cases committed at time t and prosecuted at time t ($y_{it,t}$) or at time $t+1$ ($y_{it,t+1}$). By controlling for crimes prosecuted within less than two years, we control for any possible (short-term) effect of direct elections on enforcement. Table A1 shows the results: the coefficient estimates are consistent with previous findings and even more precisely estimated.

6 Conclusions

In this paper we asked whether direct elections of the local government have affected local corruption. In order to answer this question we exploited the gradual introduction of local elections in Indonesian districts and a novel database on corruption prosecutions. We used the number of corruption crimes *committed* at a given time in a district as a measure of local corruption. Coefficient estimates are robust across various specifications and suggest that direct elections increased local corruption by about as much as the pre-election average. In order to verify whether the baseline results are driven by a possible increase in law enforcement, we estimated the impact of direct elections joint with a factor that is unambiguously associated with greater corruption but not with greater law enforcement. In addition, we estimated the impact of direct elections on corruption crimes prosecuted (rather than committed) at a given time. Both robustness checks strongly support the view that law enforcement is not driving our results.

The paper contributes to the literature on corruption by shedding new light on its institutional determinants and informing central governments about the potential costs of a form of political decentralization with direct election of the local government relative to political decentralization with indirect election of the local government.

References

Baliga, S., and T. Sjöström (1998): "Decentralization and Collusion," *Journal of Economic Theory*, 83(2), 196–232.

Banerjee, A., S. Mullainathan, and R. Hanna (2012): "Corruption," *National Bureau of Economic Research Working Paper Series*, No. 17968.

Besley, T. (2005): "Political Selection," *Journal of Economic Perspectives*, 19(3), 43–60.

Besley, T., and R. Burgess (2002): "The Political Economy of Government Responsiveness: Theory and Evidence from India," *The Quarterly Journal of Economics*, 117(4), 1415–1451.

Besley, T., and A. Case (1995): "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits," *The Quarterly Journal of Economics*, 110(3), 769–798.

Besley, T., and M. Smart (2007): "Fiscal restraints and voter welfare," *Journal of Public Economics*, 91(3âĂŞ4), 755–773.

Björkman, M., and J. Svensson (2009): "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda," *The Quarterly Journal of Economics*, 124(2), 735–769.

Burgess, R., M. Hansen, B. Olken, P. Potapov, and S. Sieber (2012): "The Political Economy of Deforestation in the Tropics," *The Quarterly Journal of Economics*, 127(4).

Dal Boť, E., and M. A. Rossi (2011): "Term Length and the Effort of Politicians," *The Review of Economic Studies*, 78(4), 1237–1263.

Ferraz, C., and F. Finan (2008): "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes," *The Quarterly Journal of Economics*, 123(2), 703–745.

——— (2011): "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments," *American Economic Review*, 101(4), 1274–1311.

Fisman, R., and R. Gatti (2002a): "Decentralization and corruption: evidence across countries," *Journal of Public Economics*, 83(3), 325–345.

——— (2002b): "Decentralization and Corruption: Evidence from U.S. Federal Transfer Programs," *Public Choice*, 113(1), 25–35.

Fukumoto, K., and Y. Horiuchi (2011): "Making Outsiders' Votes Count: Detecting Electoral Fraud through a Natural Experiment," *American Political Science Review*, 105(03), 586–603.

Gadenne, L. (2010): "Tax Me but Spend Wisely, the Political Economy of Taxes. Evidence from Brazilian Local Governments," .

Gan, L., L. C. Xu, and Y. Yao (2007): "Local elections and consumption insurance : evidence from Chinese villages," .

General Attorney Office (2011): "Annual Report," Discussion paper.

Glaeser, E. L., and R. E. Saks (2006): "Corruption in America," *Journal of Public Economics*, 90(6âĂŞ7), 1053–1072.

Gordon, S. C. (2009): "Assessing Partisan Bias in Federal Public Corruption Prosecutions," *American Political Science Review*, 103(04), 534–554.

Kunicova, J., and R. Rose-Ackerman (2001): "Electoral Rules as Constraints on Corruption: the Risks of Closed-List Proportional Representation," .

Luo, R., L. Zhang, J. Huang, and S. Rozelle (2007): "Elections, fiscal reform and public goods provision in rural China," *Journal of Comparative Economics*, 35(3), 583–611.

Martinez-Bravo, M., G. P. i. Miquel, N. Qian, and Y. Yao (2012): "The Effects of Democratization on Public Goods and Redistribution: Evidence from China," .

Mookherjee, D. (2006): "Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective," *Journal of Economic Literature*, 44(2), 367–390.

Mookherjee, D., and M. Tsumagari (2004): "The Organization of Supplier Networks: Effects of Delegation and Intermediation," *Econometrica*, 72(4), 1179–1219.

Olken, B. A. (2009): "Corruption perceptions vs. corruption reality," *Journal of Public Economics*, 93(7âĂŞ8), 950–964.

Olken, B. A., and R. Pande (2012): "Corruption in Developing Countries," *Annual Review of Economics*, 4(1), null.

Olsson, O., and M. Valsecchi (2012): "Resource Windfall and Public Goods: Evidence from a Policy Reform," .

Persson, T., G. Roland, and G. Tabellini (1997): "Separation of Powers and Political Accountability," *The Quarterly Journal of Economics*, 112(4), 1163–1202.

Persson, T., and G. Tabellini (2004): "Constitutions and Economic Policy," *Journal of Economic Perspectives*, 18(1), 75–98.

Persson, T., G. Tabellini, and F. Trebbi (2003): "ELECTORAL RULES AND CORRUPTION," *Journal of the European Economic Association*, 1(4), 958–989.

Reinikka, R., and J. Svensson (2004): "Local Capture: Evidence from a Central Government Transfer Program in Uganda," *The Quarterly Journal of Economics*, 119(2), 679–705.

——— (2011): "The power of information in public services: Evidence from education in Uganda," *Journal of Public Economics*, 95(7âĂŞ8), 956–966.

Shen, Y., and Y. Yao (2008): "Does grassroots democracy reduce income inequality in China?," *Journal of Public Economics*, 92(10âĂŞ11), 2182–2198.

Skoufias, E., A. Narayan, B. Dasgupta, and K. Kaiser (2011): "Electoral accountability, fiscal decentralization and service delivery in Indonesia," *The World Bank*, 5614.

Svensson, J. (2005): "Eight Questions about Corruption," *Journal of Economic Perspectives*, 19(3), 19–42.

Zhang, X., S. Fan, L. Zhang, and J. Huang (2004): "Local governance and public goods provision in rural China," *Journal of Public Economics*, 88(12), 2857–2871.

Zitzewitz, E. (Forthcoming): "Forensic Economics," *Journal of Economic Literature*.

Table 1

DESCRIPTIVE STATISTICS

PANEL A: TYPE OF CORRUPTION AND SECTOR DISTRIBUTION

	Total No	Total Col %	Private No	Private Col %	Public No	Public Col %	Public sphere, sector shares: admin	agr	dev	edu	food/ health	infrastructure	private	regional company	state company	missing	Total
All	1396	100.0	133	100.0	1263	100.0	54.4	4.0	2.7	11.7	5.5	5.5	2.9	1.4	1.5	10.6	100
Embezzlement	1056	75.6	35	26.3	1021	80.8	55.8	3.3	2.8	10.7	5.2	5.3	1.3	1.6	1.8	12.2	100
Conflict of interest	168	12.0	2	1.5	166	13.1	29.5	11.4	2.4	19.9	9	15.7	9.6	0.6	0.6	1.2	100
Fraud	50	3.6	7	5.3	43	3.4	67.4	4.7	0	11.6	2.3	0	9.3	0	0	4.7	100
Distribution	34	2.4	4	3.0	30	2.4	36.7	10	0	3.3	46.7	0	0	0	3.3	0	100
Extortion	20	1.4	1	0.8	19	1.5	84.2	0	0	5.3	0	0	0	0	0	10.5	100
Bribery	15	1.1	2	1.5	13	1.0	53.8	0	0	30.8	0	7.7	7.7	0	0	0	100
Illegal practice	63	4.5	51	38.3	12	1.0	75	0	8.3	0	0	0	8.3	0	0	8.3	100
Hazard	36	2.6	35	26.3	1	0.1	0	0	0	0	0	0	100	0	0	0	100

PANEL B: DISTRIBUTION OVER TIME OF CORRUPTION CASES

Distribution of corruption events: Year	Total No (1)	Total Col % (2)	Private No (3)	Private Col % (4)	Public No (5)	Public Col % (6)	Distribution of corruption crimes concerning the public sphere across districts: Year	Commited: Number of events (mean)	Commited: At least one event (mean)	Prosecuted: Number of events (mean)	Prosecuted: At least one event (mean)
1998	2	0.1	0	0.0	2	0.2	1998	0.01	0.01	0.00	0.00
1999	8	0.6	0	0.0	8	0.6	1999	0.03	0.02	0.00	0.00
2000	12	0.9	0	0.0	12	1.0	2000	0.04	0.03	0.00	0.00
2001	30	2.1	0	0.0	30	2.4	2001	0.10	0.05	0.00	0.00
2002	63	4.5	0	0.0	63	5.0	2002	0.16	0.07	0.00	0.00
2003	94	6.7	2	1.5	92	7.3	2003	0.22	0.09	0.00	0.00
2004	93	6.7	3	2.3	90	7.1	2004	0.23	0.09	0.01	0.01
2005	111	8.0	0	0.0	111	8.8	2005	0.29	0.12	0.03	0.01
2006	147	10.5	3	2.3	144	11.4	2006	0.34	0.19	0.05	0.04
2007	232	16.6	7	5.3	225	17.8	2007	0.55	0.26	0.07	0.04
2008	274	19.6	16	12.0	258	20.4	2008	0.66	0.30	0.33	0.17
2009	211	15.1	42	31.6	169	13.4	2009	0.44	0.24	0.66	0.31
2010	86	6.2	30	22.6	56	4.4	2010	0.15	0.09	0.77	0.31
2011	33	2.4	30	22.6	3	0.2	2011	0.01	0.01	0.25	0.12
Total	1396	100.0	133	100.0	1263	100.0					

Table 2

EXOGENEITY TEST

PANEL A

	Urban village	Agricultural village	%HHs in agriculture	Village area	% rice area	Population	Population per hectare	Electricity in village	Households with electricity	Slum	Households in slums	TV reception
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
Direct Elections	0.010	-0.009	-0.998	-259.281*	-0.552	8 354.488	0.974	-0.002	0.001	0.006	0.000	0.051
	(1.193)	(-1.230)	(-1.360)	(-1.740)	(-0.044)	(0.255)	(1.028)	(-0.640)	(0.169)	(1.344)	(0.621)	(0.291)
N	51 009	51 119	51 119	51 119	51 119	46 738	51 119	51 119	51 119	51 119	51 119	51 119
Adjusted R2	0.001	0.001	0.003	0.001	-0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000
outcome: mean	0.21	0.87	71.34	2419.30	224.31	817097.90	14.40	0.97	0.61	0.09	0.01	5.54
outcome: s.d.	0.40	0.34	28.36	10663.91	1566.11	597099.10	38.97	0.18	0.31	0.28	0.06	4.52

PANEL B

	Primary school	Junior-high school	Senior-high school	Hospital	Maternity house	Health center	Mosques[1]	Prayer houses[1]	Churches[1]	Buddhist temples[1]	Local newspaper	Oil and Gas transfers per capita[2]
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
Direct Elections	0.001	-0.000	0.001	-0.000	-0.000	0.002	-0.380	0.306	0.304	0.001	-0.008	-0.020*
	(0.494)	(-0.077)	(0.178)	(-0.027)	(-0.115)	(0.644)	(-1.368)	(0.341)	(0.869)	(0.068)	(-0.781)	(-1.892)
N	51 119	51 119	51 119	51 119	51 119	51 119	51 119	51 119	51 119	51 119	46 738	44 726
Adjusted R2	0.000	-0.000	-0.000	-0.000	-0.000	0.000	0.002	0.000	0.001	-0.000	0.002	0.010
outcome: mean	0.96	0.37	0.18	0.02	0.06	0.13	12.54	23.81	4.87	0.12	0.46	0.06
outcome: s.d.	0.19	0.48	0.38	0.15	0.24	0.33	11.62	26.89	14.90	1.59	0.23	0.30

PANEL C

	Traffic through land	Paved road	Distance to sub-district capital	Distance to district capital	Conflict episode (last year)	Safety post	Police post	Village head: age	Village head: male	Village head: high school	Village head: diploma	Village head: bachelor
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
Direct Elections	0.000	0.004	-0.162	-0.621	0.001	0.005	-0.002	-0.003***	0.064	0.001	-0.003	-0.003
	(0.002)	(0.397)	(-0.570)	(-0.795)	(0.683)	(0.677)	(-0.673)	(-2.623)	(0.691)	(0.184)	(-0.498)	(-0.522)
N	51 119	49 305	51 119	51 119	51 119	51 119	51 119	50 444	50 444	50 445	50 445	50 445
Adjusted R2	-0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	-0.000	0.000	0.000
outcome: mean	0.96	0.62	8.99	34.82	0.02	0.88	0.12	0.97	44.70	0.69	0.20	0.16
outcome: s.d.	0.19	0.48	13.69	30.00	0.15	0.33	0.32	0.17	8.49	0.46	0.40	0.36

Note: *** p<0.01, ** p<0.05, * p<0.1. The unit of analysis is the village. Standard errors are clustered at the district-level in brackets. Sample restricted to district which did not form or split during the period 2003-2009. Data source: PODES 2006 for all outcomes but local newspapers (PODES 2003) and resource windfall (Ministry of Finance).

[1] Number of temples for every 10,000 people.

[2] The measurement unit is millions IDR, where 1 million IDR is approximately 100 USD.

TABLE 3

BASELINE RESULTS, ONLY DISTRICTS WITH ELECTIONS IN 2005 OR AFTER 2008

DEP VARIABLE	At least one corruption event					Number of corruption events				
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A: ALL CORRUPTION CRIMES										
DIRECT ELECTIONS	0.133***	0.081**	0.096***	0.092***	0.069*	0.274***	0.152	0.215*	0.205*	0.151
	(0.026)	(0.033)	(0.034)	(0.034)	(0.039)	(0.090)	(0.107)	(0.114)	(0.115)	(0.112)
R-squared	0.030	0.049	0.076	0.086	0.179	0.015	0.027	0.042	0.058	0.223
PANEL B: EMBEZZLEMENT										
DIRECT ELECTIONS	0.112***	0.064**	0.081**	0.076**	0.057	0.230***	0.123	0.184*	0.173	0.120
	(0.025)	(0.031)	(0.032)	(0.032)	(0.038)	(0.083)	(0.100)	(0.107)	(0.109)	(0.103)
R-squared	0.024	0.042	0.069	0.079	0.164	0.012	0.023	0.036	0.053	0.218
PANEL C: CONFLICT OF INTEREST										
DIRECT ELECTIONS	0.035***	0.030**	0.028*	0.029**	0.021	0.051**	0.047**	0.049**	0.052**	0.038*
	(0.012)	(0.015)	(0.015)	(0.015)	(0.014)	(0.020)	(0.022)	(0.023)	(0.025)	(0.022)
R-squared	0.012	0.018	0.020	0.034	0.143	0.010	0.016	0.019	0.030	0.131
PANEL D: CONFLICT OF INTEREST, BRIBERY AND EXTORTION										
DIRECT ELECTIONS	0.037***	0.032**	0.031**	0.032**	0.025*	0.037***	0.032**	0.031**	0.032**	0.025*
	(0.013)	(0.015)	(0.015)	(0.015)	(0.015)	(0.013)	(0.015)	(0.015)	(0.015)	(0.015)
R-squared	0.013	0.019	0.021	0.034	0.161	0.013	0.019	0.021	0.034	0.161
Time Effects		year	year	region year	province year		year	year	region year	province year
Fixed Effects			district	district	district			district	district	district
Observations	2,112	2,112	2,112	2,112	2,112	2,112	2,112	2,112	2,112	2,112
Number of clusters	214	214	214	214	214	214	214	214	214	214
Number of districts	268	268	268	268	268	268	268	268	268	268

*** p<0.01, ** p<0.05, * p<0.1 Standard errors (in brackets) clustered at the district level, using district borders as in 1999.

TABLE 4

BASELINE RESULTS, ALL DISTRICTS

DEP VARIABLE	At least one corruption event					Number of corruption events				
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A: ALL CORRUPTION CRIMES										
DIRECT ELECTIONS	0.187***	0.062*	0.056*	0.052	0.039	0.406***	0.145	0.161*	0.153	0.125
	(0.021)	(0.035)	(0.032)	(0.033)	(0.034)	(0.067)	(0.106)	(0.093)	(0.094)	(0.086)
R-squared	0.063	0.084	0.124	0.134	0.215	0.037	0.049	0.075	0.090	0.181
PANEL B: EMBEZZLEMENT										
DIRECT ELECTIONS	0.161***	0.065*	0.067**	0.061*	0.053	0.308***	0.118	0.141*	0.137	0.104
	(0.020)	(0.034)	(0.031)	(0.032)	(0.034)	(0.059)	(0.095)	(0.083)	(0.085)	(0.080)
R-squared	0.052	0.068	0.106	0.120	0.193	0.027	0.037	0.059	0.078	0.161
PANEL C: CONFLICT OF INTEREST										
DIRECT ELECTIONS	0.050***	0.020	0.019	0.015	0.008	0.082***	0.046*	0.047*	0.044	0.035
	(0.009)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.026)	(0.027)	(0.028)	(0.030)
R-squared	0.021	0.026	0.030	0.043	0.128	0.020	0.024	0.028	0.036	0.110
PANEL D: CONFLICT OF INTEREST, BRIBERY AND EXTORTION										
DIRECT ELECTIONS	0.056***	0.023	0.018	0.014	0.011	0.057***	0.023	0.019	0.014	0.011
	(0.010)	(0.017)	(0.016)	(0.016)	(0.017)	(0.010)	(0.017)	(0.016)	(0.017)	(0.017)
R-squared	0.024	0.030	0.034	0.046	0.143	0.023	0.029	0.034	0.045	0.135
Pre-election dummy	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects		year	year	region year	province year		year	year	region year	province year
Fixed Effects			district	district	district			district	district	district
Observations	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470
Number of clusters	255	255	255	255	255	255	255	255	255	255
Number of districts	349	349	349	349	349	349	349	349	349	349

*** p<0.01, ** p<0.05, * p<0.1 Standard errors (in brackets) clustered at the district level, using district borders as in 1999.

TABLE 5

BASELINE RESULTS, ALL DISTRICTS

DEP VARIABLE	At least one corruption event					Number of corruption events				
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A: ALL CORRUPTION CRIMES										
ELECTION YEAR	0.148***	0.060*	0.057*	0.051	0.041	0.366***	0.166	0.173*	0.168*	0.135
	(0.022)	(0.032)	(0.031)	(0.032)	(0.034)	(0.075)	(0.102)	(0.094)	(0.099)	(0.098)
FOLLOWING YEARS	0.208***	0.065	0.055	0.052	0.034	0.428***	0.124	0.130	0.112	0.100
	(0.025)	(0.045)	(0.043)	(0.043)	(0.047)	(0.078)	(0.128)	(0.120)	(0.115)	(0.103)
R-squared	0.065	0.084	0.124	0.134	0.215	0.037	0.049	0.075	0.090	0.181
PANEL B: EMBEZZLEMENT										
ELECTION YEAR	0.132***	0.065**	0.067**	0.061*	0.054	0.275***	0.123	0.141*	0.142	0.106
	(0.022)	(0.031)	(0.031)	(0.031)	(0.034)	(0.068)	(0.094)	(0.085)	(0.090)	(0.091)
FOLLOWING YEARS	0.177***	0.065	0.065	0.061	0.051	0.326***	0.114	0.140	0.125	0.101
	(0.024)	(0.043)	(0.041)	(0.041)	(0.044)	(0.067)	(0.110)	(0.104)	(0.100)	(0.088)
R-squared	0.053	0.068	0.106	0.120	0.193	0.028	0.037	0.059	0.078	0.161
PANEL C: CONFLICT OF INTEREST										
ELECTION YEAR	0.051***	0.030*	0.026	0.021	0.013	0.088***	0.067**	0.060*	0.058*	0.048
	(0.013)	(0.017)	(0.017)	(0.018)	(0.018)	(0.025)	(0.031)	(0.032)	(0.033)	(0.035)
FOLLOWING YEARS	0.049***	0.010	0.002	-0.002	-0.004	0.079***	0.026	0.012	0.010	0.002
	(0.010)	(0.020)	(0.019)	(0.019)	(0.020)	(0.018)	(0.033)	(0.032)	(0.033)	(0.032)
R-squared	0.021	0.027	0.031	0.044	0.129	0.020	0.025	0.030	0.038	0.111
PANEL D: CONFLICT OF INTEREST, BRIBERY AND EXTORTION										
ELECTION YEAR	0.050***	0.028	0.023	0.018	0.013	0.053***	0.029*	0.024	0.018	0.013
	(0.013)	(0.017)	(0.018)	(0.018)	(0.018)	(0.014)	(0.017)	(0.018)	(0.019)	(0.019)
FOLLOWING YEARS	0.059***	0.017	0.007	0.004	0.005	0.060***	0.016	0.006	0.002	0.005
	(0.012)	(0.021)	(0.020)	(0.020)	(0.021)	(0.013)	(0.022)	(0.020)	(0.020)	(0.021)
R-squared	0.024	0.030	0.035	0.046	0.143	0.023	0.030	0.034	0.046	0.135
Pre-election dummy	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects		year	year	region year	province year		year	year	region year	province year
Fixed Effects			district	district	district			district	district	district
Observations	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470
Number of clusters	255	255	255	255	255	255	255	255	255	255
Number of districts	349	349	349	349	349	349	349	349	349	349

*** p<0.01, ** p<0.05, * p<0.1 Standard errors (in brackets) clustered at the district level, using district borders as in 1999.

TABLE 6

ROBUSTNESS CHECK: EXCLUDE ONE REGION AT A TIME

DEP VARIABLE	At least one corruption event					Number of corruption events				
Excluded region	Jawa	Kalimantan	Nusa Tenggara & Maluku	Sulawesi	Sumatera	Jawa	Kalimantan	Nusa Tenggara & Maluku	Sulawesi	Sumatera
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A: ALL CORRUPTION CRIMES										
DIRECT ELECTIONS	0.046	0.066*	0.036	0.079**	0.050	0.223*	0.152	0.122	0.207**	0.109
	(0.037)	(0.035)	(0.034)	(0.034)	(0.039)	(0.120)	(0.095)	(0.099)	(0.104)	(0.104)
R-squared	0.106	0.132	0.123	0.132	0.125	0.053	0.109	0.073	0.079	0.070
PANEL B: ALL CORRUPTION CRIMES										
ELECTION YEAR	0.057	0.060*	0.039	0.076**	0.047	0.255**	0.143	0.141	0.219**	0.116
	(0.036)	(0.034)	(0.033)	(0.033)	(0.037)	(0.125)	(0.089)	(0.101)	(0.106)	(0.111)
FOLLOWING YEARS	0.015	0.081*	0.028	0.085*	0.058	0.133	0.174	0.073	0.176	0.092
	(0.049)	(0.047)	(0.046)	(0.046)	(0.052)	(0.148)	(0.131)	(0.130)	(0.136)	(0.123)
R-squared	0.108	0.132	0.123	0.132	0.125	0.054	0.110	0.073	0.079	0.070
PANEL C: EMBEZZLEMENT										
DIRECT ELECTIONS	0.067*	0.079**	0.049	0.072**	0.064*	0.176*	0.143*	0.112	0.159*	0.120
	(0.035)	(0.035)	(0.033)	(0.034)	(0.037)	(0.106)	(0.083)	(0.090)	(0.095)	(0.097)
R-squared	0.085	0.116	0.104	0.117	0.106	0.037	0.098	0.057	0.064	0.053
PANEL D: EMBEZZLEMENT										
ELECTION YEAR	0.076**	0.077**	0.051	0.070**	0.060	0.187*	0.135*	0.116	0.157	0.118
	(0.035)	(0.034)	(0.033)	(0.034)	(0.037)	(0.110)	(0.079)	(0.092)	(0.097)	(0.105)
FOLLOWING YEARS	0.041	0.085*	0.045	0.078*	0.072	0.146	0.164	0.100	0.164	0.127
	(0.045)	(0.045)	(0.044)	(0.045)	(0.049)	(0.131)	(0.113)	(0.113)	(0.118)	(0.106)
R-squared	0.086	0.116	0.104	0.117	0.106	0.037	0.098	0.057	0.064	0.053
Pre-election dummy	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects	year	year	year	year	year	year	year	year	year	year
Fixed Effects	district	district	district	district	district	district	district	district	district	district
Observations	2,220	2,998	3,178	2,956	2,528	2,220	2,998	3,178	2,956	2,528
Number of clusters	234	300	317	295	250	234	300	317	295	250
Number of districts	145	227	234	218	196	145	227	234	218	196

*** p<0.01, ** p<0.05, * p<0.1 Standard errors (in brackets) clustered at the district level, using district borders as in 1999.

TABLE 7

INTERACTION WITH PER CAPITA OIL AND GAS TRANSFERS

DEP VARIABLE	At least one corruption event					Number of corruption events				
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A: ALL CORRUPTION CRIMES										
DIRECT ELECTIONS	0.220***	0.048	0.049	0.050	0.029	0.507***	0.161	0.164	0.166	0.111
	(0.023)	(0.036)	(0.036)	(0.036)	(0.038)	(0.081)	(0.106)	(0.105)	(0.103)	(0.095)
R-squared	0.078	0.102	0.136	0.148	0.239	0.050	0.064	0.087	0.104	0.206
PANEL B: ALL CORRUPTION CRIMES										
DIRECT ELECTIONS	0.221***	0.037	0.040	0.040	0.018	0.474***	0.033	0.037	0.035	-0.030
	(0.024)	(0.037)	(0.037)	(0.037)	(0.039)	(0.074)	(0.098)	(0.100)	(0.097)	(0.090)
ELECTIONS × WINDFALL	0.217*	0.300***	0.248**	0.255***	0.288***	3.338*	3.533**	3.515**	3.564**	3.832**
	(0.126)	(0.109)	(0.096)	(0.094)	(0.101)	(1.702)	(1.649)	(1.717)	(1.700)	(1.628)
R-squared	0.081	0.106	0.141	0.154	0.242	0.108	0.127	0.159	0.176	0.276
PANEL C: EMBEZZLEMENT										
DIRECT ELECTIONS	0.199***	0.049	0.052	0.051	0.036	0.365***	0.014	0.018	0.019	-0.045
	(0.023)	(0.036)	(0.036)	(0.036)	(0.039)	(0.065)	(0.089)	(0.090)	(0.088)	(0.085)
ELECTIONS × WINDFALL	0.211*	0.279**	0.208**	0.214**	0.229*	3.331*	3.491**	3.438*	3.469**	3.715**
	(0.124)	(0.109)	(0.102)	(0.102)	(0.119)	(1.705)	(1.665)	(1.748)	(1.741)	(1.671)
R-squared	0.073	0.095	0.123	0.138	0.221	0.118	0.134	0.162	0.181	0.268
PANEL D: CONFLICT OF INTEREST										
DIRECT ELECTIONS	0.055***	0.015	0.015	0.012	0.004	0.091***	0.049	0.048	0.047	0.036
	(0.010)	(0.019)	(0.018)	(0.018)	(0.018)	(0.019)	(0.032)	(0.032)	(0.033)	(0.033)
ELECTIONS × WINDFALL	-0.011	0.007	0.032	0.042	0.064**	-0.025	-0.007	0.044	0.063	0.070
	(0.016)	(0.020)	(0.024)	(0.027)	(0.031)	(0.020)	(0.026)	(0.044)	(0.056)	(0.053)
R-squared	0.026	0.031	0.036	0.062	0.155	0.023	0.027	0.032	0.048	0.127
PANEL E: CONFLICT OF INTEREST, BRIBERY AND EXTORTION										
DIRECT ELECTIONS	0.061***	0.017	0.016	0.014	0.009	0.064***	0.017	0.017	0.014	0.009
	(0.010)	(0.019)	(0.018)	(0.018)	(0.019)	(0.011)	(0.019)	(0.019)	(0.019)	(0.019)
ELECTIONS × WINDFALL	-0.013	0.008	0.033	0.042	0.063**	-0.015	0.007	0.033	0.042	0.062**
	(0.016)	(0.020)	(0.024)	(0.026)	(0.031)	(0.016)	(0.020)	(0.024)	(0.026)	(0.031)
R-squared	0.028	0.034	0.040	0.061	0.162	0.026	0.033	0.039	0.059	0.151
Pre-election dummy	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Resource × year	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects		year	year	region year	province year		year	year	region year	province year
Fixed Effects			district	district	district			district	district	district
Observations	2,988	2,988	2,988	2,988	2,988	2,988	2,988	2,988	2,988	2,988
Number of clusters	246	246	246	246	246	246	246	246	246	246
Number of districts	275	275	275	275	275	275	275	275	275	275

*** p<0.01, ** p<0.05, * p<0.1 Standard errors (in brackets) clustered at the district level, using district borders as in 1999.

TABLE 8

FALSIFICATION EXPERIMENT

DEP VARIABLE	At least one corruption event					Number of corruption events				
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A: ALL CORRUPTION CRIMES										
DIRECT ELECTIONS	0.089***	0.005	0.005	0.001	-0.001	0.163***	0.002	0.004	-0.005	-0.002
	(0.011)	(0.013)	(0.014)	(0.014)	(0.016)	(0.025)	(0.032)	(0.030)	(0.028)	(0.036)
R-squared	0.049	0.095	0.109	0.128	0.228	0.033	0.068	0.078	0.096	0.205
PANEL B: ALL CORRUPTION CRIMES										
ELECTION YEAR	0.068***	0.005	0.005	-0.001	-0.003	0.135***	0.007	0.007	-0.006	-0.001
	(0.015)	(0.017)	(0.017)	(0.016)	(0.020)	(0.039)	(0.043)	(0.041)	(0.038)	(0.046)
FOLLOWING YEARS	0.103***	0.004	0.005	0.007	0.002	0.182***	-0.010	-0.004	-0.005	-0.007
	(0.013)	(0.017)	(0.018)	(0.018)	(0.019)	(0.027)	(0.038)	(0.038)	(0.037)	(0.034)
R-squared	0.052	0.095	0.109	0.128	0.228	0.034	0.068	0.078	0.096	0.205
PANEL C: EMBEZZLEMENT										
DIRECT ELECTIONS	0.079***	0.008	0.009	0.005	0.004	0.122***	0.003	0.007	0.001	0.005
	(0.010)	(0.013)	(0.013)	(0.013)	(0.015)	(0.020)	(0.030)	(0.028)	(0.025)	(0.033)
R-squared	0.043	0.078	0.090	0.110	0.211	0.027	0.052	0.059	0.076	0.175
PANEL D: EMBEZZLEMENT										
ELECTION YEAR	0.065***	0.010	0.011	0.005	0.005	0.108***	0.010	0.011	0.003	0.010
	(0.014)	(0.016)	(0.016)	(0.015)	(0.019)	(0.031)	(0.039)	(0.036)	(0.033)	(0.042)
FOLLOWING YEARS	0.088***	0.002	0.005	0.005	0.001	0.131***	-0.013	-0.005	-0.004	-0.009
	(0.012)	(0.016)	(0.017)	(0.017)	(0.018)	(0.022)	(0.032)	(0.031)	(0.030)	(0.027)
R-squared	0.044	0.078	0.090	0.110	0.211	0.027	0.052	0.059	0.076	0.175
Pre-election dummy	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects		year	year	region year	province year		year	year	region year	province year
Fixed Effects			district	district	district			district	district	district
Observations	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470
Number of clusters	255	255	255	255	255	255	255	255	255	255
Number of districts	349	349	349	349	349	349	349	349	349	349

*** p<0.01, ** p<0.05, * p<0.1 Standard errors (in brackets) clustered at the district level, using district borders as in 1999.

TABLE A1

CONTROL FOR CORRUPTION CRIMES COMMITED AT TIME T AND PROSECUTED AT TIME T AND T+1

DEP VARIABLE	At least one corruption event					Number of corruption events				
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL A: ALL CORRUPTION CRIMES										
DIRECT ELECTIONS	0.115***	0.059*	0.049*	0.050*	0.038	0.228***	0.180**	0.170**	0.169**	0.144*
	(0.019)	(0.031)	(0.029)	(0.029)	(0.031)	(0.059)	(0.089)	(0.083)	(0.083)	(0.077)
R-squared	0.265	0.275	0.299	0.306	0.366	0.274	0.279	0.289	0.299	0.352
PANEL B: ALL CORRUPTION CRIMES										
ELECTION YEAR	0.087***	0.049*	0.046*	0.047	0.036	0.217***	0.181*	0.177**	0.181**	0.152*
	(0.020)	(0.029)	(0.028)	(0.029)	(0.030)	(0.074)	(0.095)	(0.087)	(0.091)	(0.089)
FOLLOWING YEARS	0.130***	0.069*	0.057	0.058	0.041	0.235***	0.178*	0.151	0.140	0.124
	(0.021)	(0.038)	(0.038)	(0.038)	(0.040)	(0.067)	(0.100)	(0.101)	(0.096)	(0.086)
R-squared	0.266	0.275	0.299	0.306	0.366	0.274	0.279	0.289	0.299	0.352
PANEL C: EMBEZZLEMENT										
DIRECT ELECTIONS	0.100***	0.058*	0.056**	0.056*	0.048	0.181***	0.149*	0.159**	0.161**	0.130*
	(0.018)	(0.031)	(0.028)	(0.029)	(0.030)	(0.050)	(0.080)	(0.074)	(0.075)	(0.069)
R-squared	0.253	0.261	0.285	0.293	0.347	0.251	0.255	0.269	0.281	0.329
PANEL D: EMBEZZLEMENT										
ELECTION YEAR	0.083***	0.056**	0.056**	0.057**	0.049	0.190***	0.162*	0.168**	0.175**	0.141*
	(0.019)	(0.028)	(0.028)	(0.028)	(0.030)	(0.066)	(0.086)	(0.078)	(0.082)	(0.079)
FOLLOWING YEARS	0.109***	0.059	0.055	0.055	0.045	0.176***	0.137	0.136	0.126	0.103
	(0.021)	(0.037)	(0.037)	(0.037)	(0.039)	(0.056)	(0.089)	(0.089)	(0.086)	(0.077)
R-squared	0.253	0.261	0.285	0.293	0.347	0.251	0.256	0.269	0.281	0.329
Pre-election dummy	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Cases prosecuted at t and t+1	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Time Effects		year	year	region year	province year		year	year	region year	province year
Fixed Effects			district	district	district			district	district	district
Observations	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470	3,470
Number of clusters	255	255	255	255	255	255	255	255	255	255
Number of districts	349	349	349	349	349	349	349	349	349	349

*** p<0.01, ** p<0.05, * p<0.1 Standard errors (in brackets) clustered at the district level, using district borders as in 1999.

Paper 4

Resource Windfalls and Public Goods: Evidence From a Policy Reform*

Ola Olsson
University of Gothenburg

Michele Valsecchi†
University of Gothenburg

October 26, 2012

Abstract

In this paper, we outline an empirical approach for understanding whether natural resource windfalls have a positive or negative impact on local governments' provision of public goods. The literature on the curse of natural resources suggests that resource windfalls might not necessarily lead to good economic outcomes and that rents might be squandered in corruption and rent seeking. In order to identify the impact of natural resources on local government behavior, we exploit a country-wide fiscal decentralization reform in Indonesia, providing producing provinces a direct share of resource revenues. Our identification strategy is to compare villages along the border of three producing provinces in Sumatra and Kalimantan before and after the legislative change. Detailed descriptive statistics on district government budgets confirm the goodness of the research design. Regression analysis on a wide range of public goods suggests that the revenue windfall had a positive impact on the prevalence of high schools and various other public goods. We find no evidence of a resource curse.

Key words: Natural Resource Curse; Oil; Fiscal Decentralization; Public Goods.

JEL Classification codes: E62,H11,H41,H71,H72,O11,O13,Q32,Q33.

*We would like to thank Giacomo de Luca, Dilip Mookherjee, Michael Callen, Alessandro Tarozzi and seminar participants at LICOS (Leuven), IV Summer Workshop in Development Economics (Ascea), and Gothenburg for useful comments. Special thanks to Måns Söderbom for many useful discussions. Michele Valsecchi thanks the Royal Society of Arts and Sciences in Gothenburg and the Knut and Alice Wallenberg Foundation for financial support.

†Corresponding author. Email: michele.valsecchi@economics.gu.se.

1 Introduction

Several developing countries around the world are currently enjoying a strong boom in natural resource revenues. Very high world market prices of oil and minerals have resulted in abnormal growth rates in some very poor countries and to an intensive prospecting activity even in countries that were previously not extracting such resources.

At the same time, there is a widespread awareness that resource rents do not necessarily provide a foundation for sustainable economic development. In the very extensive literature on the *curse of natural resources*, it has been shown that countries with substantial natural resource rents often have had a relatively weak economic performance compared to resource-poor countries. Several intermediate channels for this adverse impact have been proposed; Dutch disease effects from currency appreciation, a crowding out of education and innovation, and a higher degree of rent seeking and corruption.[1] In particular, the political economy of the resource curse has received a lot of attention during recent years.[2] However, this mainly theoretical or macro-oriented literature has so far not been able to document many cases of how resource windfalls have actually been used in the public sector. Given this shortcoming, is has been difficult to provide evidence-based policy advice to governments that are concerned about being affected by the resource-curse.

In this paper, we analyze the impact of a country-wide reform from 2001 that decentralized Indonesia's resource revenues (oil and gas) from the federal government to provinces. Our analysis investigates the effect of these windfall gains on local public goods provision in two resource-rich regions; Sumatra and the Indonesian part of Borneo, Kalimantan. The reform allowed resource producing provinces to obtain a percentage of the natural resource revenues collected by the central government. The objective of our study is to determine whether this policy reform actually led to more local public goods such as schools, health clinics, and infrastructure. The main hypothesis is that local public goods should have increased. However, evidence on the resource curse from other countries would rather suggest no impact

[1] See for instance Sachs and Warner (1997), Sachs and Warner (2001), and Gylfason (2001). van der Ploeg (2011) supplies a recent overview of the literature.

[2] See for instance Torvik (2002) and Robinson, Torvik, and Verdier (2006).

or perhaps even a negative impact.

The main contribution of our paper is the research design that we propose for studying the question at hand. Our identification strategy in the empirical analysis is to compare public good outcomes in villages located in producing provinces with neighboring villages across the border in non-producing provinces, before and after the reform. The first category makes up our treatment group (that obtained resource revenues) whereas the second category is our control group (that received no resource revenue). We argue that this design makes our study close to a natural experiment. Our data allows us to study public goods outcomes during the period 2001-2005. We employ a regression discontinuity design with the border between the producing and non-producing provinces as our forcing variable. The descriptive statistics as well as a broad general analysis of the characteristics of the provinces suggest that no key differences exist between treatment and control areas apart from the fact that the treated districts receive resource revenues.

The results are mixed. In general, there are no indications of a decrease in public goods as a result of the reform. In this sense, we find no evidence of a resource curse. Our results suggest that high school facilities tended to improve in treated villages, in particular 3-4 years after the decentralization. We also find evidence that other public goods improve, but they vary across the two study areas. For other public goods like access to clean water, the reform appears to have no impact on either Sumatra or Kalimantan.

Our results are related to a small but growing literature on the impact of resource windfalls within countries. Apart from Indonesia, Brazil has also chosen to decentralize oil revenues to producing areas. Using data on Brazilian municipalities, Monteiro and Ferraz (2009) show that oil revenue appears to have led to an increase in the number of public employees but not to the provision of public goods like health and education. Their main focus is on political economy aspects and their analysis demonstrates that the windfall created a large incumbency advantage in local elections. In their empirical analysis, the authors further analyze how oil revenues affected policy outcomes. The central result in this regard is that whereas oil windfalls increased the number of public employees, they had no significant impact on education or health supply.

Oil windfalls in Brazil is also the topic in Brollo, Nannicini, Perotti, and Tabellini (Forthcoming) and Caselli and Michaels (Forthcoming). Both recognize that municipalities with oil revenues have increased their spending, but like Monteiro and Ferraz (2009), these studies emphasize that the increased spending has not improved public goods as much as one would have expected. Caselli and Michaels (Forthcoming) further show that oil windfalls are associated with illegal activities by mayors, suggesting an increase in corrupt behaviors. Brollo, Nannicini, Perotti, and Tabellini (Forthcoming) use a regression discontinuity design and provide evidence that larger windfalls increase corruption and lower the quality of political candidates on the local level.[3]

Our paper obviously makes a contribution by having a different and, for this purpose, a novel object of study; Indonesia. Our research design is further different since we use a border between a producing and a non-producing province as our forcing variable and exploit the time variation in public goods outcomes. The methodology that we employ is most similar to Dell (2010) who also uses a border as a forcing variable in a historical analysis of the long-run legacy of a colonial institution in Peru.

Our quasi-experimental approach is further related to the large literature on randomized control trials in developing countries (see for instance Duflo, Glennerster, and Kremer (2007)). In recent research, it has often been emphasized that an up-scaling of micro field experiments is a natural direction for future work. The current paper might be seen as an attempt to contribute to this agenda. Furthermore, our approach is related to that of Reinikka and Svensson (2004) who study the extent to which a new grant from the Ugandan government actually reached 250 schools. In a similar spirit, we also combine data on grants from a central government with village level data on actual public good provision.

Our paper is also related to a large literature on the pros and cons of fiscal decentralization. Many of these works use cross-country regressions to assess whether fiscal decentralization is associated with higher economic growth.[4] Other studies use cross-regional variation within

[3] See also Vicente (2010) who compares outcomes in Sao Tome and Principe with Cape Verde as a "control country".

[4] See for instance Davoodi and Zou (1998).

large countries such as United States (Akai and Sakata (2002)) or China (Zhang and Zou (1998)). Our paper is closest to Skoufias, Narayan, Dasgupta, and Kaiser (2011) analysis of the recent reform towards direct elections in districts in Indonesia, showing that the electoral reforms had a positive impact public goods spending. Unlike their study, our treatment is resource windfalls rather than the introduction of direct elections.

The paper is structured as follows: Section 2 gives a brief theoretical background to the political economy of resource windfalls. Section 3 provides information about the context of the study and section 4 discusses the data and identification strategy. The main results are presented in section 5. Section 6 provides the details for the second study area whereas section 6 concludes.

2 The political economy of resource windfalls

From a policy point of view, there are at least four potential strategies that countries can choose regarding resource revenues. The most common strategy when it comes to oil and gas is probably to maintain a close central government control of the extraction process and then keep all the revenue at the center at the full discretion of the incumbent government. This is the path chosen by many African and Middle Eastern countries such as Sudan, Nigeria, and Saudi Arabia. This is also the policy associated with the most blatant failures. As discussed by Sala-i Martin and Subramanian (2003), Nigeria experienced a massive inflow of oil money from the 1970s, yet income per capita in the late 1990s was not higher than at independence. More or less all of the revenue disappeared inside corrupt governments. Several cross-country empirical studies have further indicated that natural resource revenues are associated with more corruption (Leite and Weidmann (1999); Dalgaard and Olsson (2008)).

A second strategy is to keep revenues on central level but to save (or "lock up") the money in a fund that cannot be used freely by incumbent governments. This strategy has been followed by Norway and seems to be appropriate in rich countries that face the risk of Dutch disease through an appreciation of their exchange rates. As argued by Collier (2008), the oil fund-strategy seems less appropriate for poor countries that are seriously constrained by poor infrastructure and low levels of health care and education. In such countries, large scale public investments are often necessary to maintain a sustainable growth process.

A third strategy, which does not seem to have ever been tried in reality, is to redistribute all resource incomes back to the households. Sala-i Martin and Subramanian (2003) suggested such a policy for Nigeria. Although such a policy would of course have great risks and great difficulties (how should such payments be distributed, for instance), the authors contend that it should at least not lead to a worse outcome than the previously followed policy in Nigeria of keeping all revenue at the discretion of the government.

A fourth alternative, which actually has been tried in at least Brazil and Indonesia, is to redistribute a substantial portion of resource revenues to the regional or local level. Local governments would typically at the same time get an increased responsibility for providing public goods. What are the potential benefits and disadvantages of this strategy?

To the extent that the public goods to be provided are truly local, a greater local autonomy over their provision should improve the matching between local preferences and the policy choices made. On the other hand, if there are obvious economies of scale involved or if the public goods are not necessarily only local (like roads that run through several districts), policy decisions should be made at a higher administrative level.

However, even if the public goods are truly local, it is not necessarily the case that a boom in resource revenues will lead to more public goods. The actual outcome will typically depend crucially on the nature of local political institutions.[5] On a macro level, it has been shown that environments with strong private property rights and more accountable governments are more likely to experience economic growth in response to resource booms (Mehlum, Moene, and Torvik (2006)). On the micro level, we have mainly the Brazilian studies to use as a benchmark. Although local governments were democratically elected in the studied municipalities, the main tendency appeared to be that the actual quality or quantity of public goods did not really increase, although spending did increase. Even in democratic settings such as Brazilian municipalities, resource windfalls might thus not lead to better economic outcomes.

[5] The analysis in Reinikka and Svensson (2004) shows for instance how local governments typically captured a substantial share of grants intended for schools in Uganda.

The country that we analyze in this study is Indonesia which experienced similar fiscal decentralization reforms as Brazil and a similar boom in oil revenues. The main research topic is whether resource windfall gains resulted in an increase in local public goods provision. Judging from existing studies, our main hypothesis is that resource windfalls should lead to more public goods, although we recognize that such an effect cannot be taken for granted due to the potentially confounding impact of local political institutions.

3 Context of study

3.1 The 1999 Fiscal Decentralization Reform

The Indonesian administrative structure is composed of different levels: central government, provinces (like US states), districts (US counties), sub-districts and villages. During the 1966-1998 autocratic regime, most of the power was retained by the central government. After the fall of Suharto, the government undertook a massive decentralization process and redistributed a large part of this power to districts. The transfer of authority concerned all fields other than macro-policies[6]: public works, health, education and culture, agriculture, transportation, industry and trade, investment, environment, land, cooperatives, and labor (art 11.2).[7] The reform became law in November 1999 and came into power simultaneously across all Indonesian districts in January 2001.

Laws 22/1999 and 25/1999 regulate the sources of local revenue. They consist of: own income (local taxes and fees, returns from regional-owned enterprises), revenue sharing (local share of taxes, local share of revenues from natural resources) and grants (transfers from the central government). The greatest part of local revenue used to come from transfers from the central government in the pre-decentralization period (called SDO) and continues to be

[6]Macro policies include foreign politics, defense, justice, monetary, fiscal and religious policies.

[7]It is difficult to find additional details on these responsibilities. About education: since 1994 education is mandatory until the 9th grade, therefore districts are particularly responsible for primary and junior-high education. It is not clear how provinces and districts share the responsibility for school building and for hiring and paying teachers. About infrastructures: districts are not directly responsible for electricity provision because that is typically provided by a State-owned enterprise (PLN); they are directly responsible for water provision because that is typically provided by local branches of the water utilities (PDAM). About roads: the central government is directly responsible for highways; provinces are directly responsible for roads crossing more than one district; districts are directly responsible for all the others.

so even after decentralization (DAU and DAK). Among the other sources of income, one was deeply affected by the reform and constitutes the focus of this paper: the redistribution of revenues from natural resources. Natural resources are oil, natural gas, mining, forestry and fishing. While state income from fishing was redistributed equally across all districts, the revenues from all the other resources were redistributed according to location. Table 1 shows the exact shares which went to central and regional governments (art. 6 of Law 25/1999).

Following decentralization, the central government retained a lower percentage of the natural resource tax revenues, while resource-abundant districts retained a greater percentage. Resource-abundant districts (henceforth: producing districts) were not the only beneficiaries of this re-allocation. The fiscal decentralization law states also that non-producing districts within producing provinces are entitled to a share of natural resource tax revenues. This share varies depending on the type of natural resource (see table 1). Although it is relatively high for forestry and mining and low for oil and gas, the latter are a lot more valuable. Therefore, this legislative change not only provides producing districts with a substantial share of the resource revenues, but also redistributes another share to districts located nearby. A noticeable feature of the revenue sharing originating from natural resources is that the law does not specify how the receiving districts should spend these additional revenues, i.e. there are no obligations attached to them.

Table 1: Allocation of revenues from natural resources (percentages)

Period	< 2001			≥ 2001			
Type	Centre	Province	Districts	Centre	Province	Districts	
		Prod.	Prod.		Prod.	Prod.	Non-prod.
Oil	100.0	0.0	0.0	84.5	3.1	6.2	6.2
Gas	100.0	0.0	0.0	69.5	6.1	12.2	12.2
Mining, rent	65.0	19.0	16.0	20.0	16.0	6.4	0.0
Mining, royalty	30.0	56.0	14.0	20.0	16.0	3.2	32.0
Forestry	55.0	30.0	15.0	20.0	16.0	3.2	32.0

Source: World Bank (1994) and Law 25/1999

3.2 Study areas: Sumatra and Kalimantan

The two areas that we study in this paper are Sumatra and the Indonesian part of Borneo (Kalimantan). There are several oil and gas producing provinces in Indonesia. However, only few of them produce a quantity of oil and gas that qualifies transfers to non-producing districts located in producing provinces greater than 5 percent of their district budget. These provinces are located in Sumatra and Kalimantan. Figures 1-2 show the distribution of oil and gas revenues on province and district level on Sumatra. As the Figures show, the most central province, called Riau, has received substantial new revenues from natural resources after the recent reform and so has the province of South Sumatra. Also the northern province, Aceh, receives large flows of rents. However, this province has for a long time sought independence from Indonesia and has been plagued by civil strife. Aceh is also one of the provinces that was hardest hit by the 2004 tsunami, just like all the provinces with shores along the Sumatran west coast.[8] In order to avoid these confounding effects, our main analysis will exclusively focus on the provinces Riau and South Sumatra as treatment regions and its neighboring districts in North Sumatra in the northwest and Jambi in the south as control regions.

Our main approach is thus to use the areas along the borders of natural resource-rich Riau and South Sumatra to identify the impact of resource-related revenues on public good provision on Sumatra. The dotted lines in Figure 3 specify more exactly the borders that we focus on in this part of the study. In the empirical section, we will use data from villages that are 200-300 km from the border on each side.

Spatial RD designs require all determinants of the outcome of interest to change *smoothly* at the border. One potential threat to this econometric methodology is that the province border was drawn in correspondence of geographic or human cleavages which are themselves correlated with the provision of public goods (see for instance Dell (2010)). Figure 3 shows the geography of our treatment and control areas. The Figure shows that there are no obvious discontinuities along the border in terms of terrain. Jambi, located in between the two treatment areas, has a very similar lowland geography to Riau and South Sumatra,

[8] The big tsunami in december 2004 had its epicenter in the Indian Ocean right west of Aceh. Hence, Aceh and the provinces on the western coast of Sumatra were affected but not the eastern coasts.

whereas North Sumatra has lowland plains close to the border but also mountains in the west. In the empirical section, we discuss further how we exclude mountain villages in order to check robustness.

As Figure 3 shows, Riau is located by the Strait of Malacca and has the Singapore and Kuala Lumpur metropolitan areas as neighbors across the strait. Riau province has currently about 5.5 million inhabitants and has experienced a steady growth of population and of its economy since the 1970s, largely due to natural resource exploitation. South Sumatra's population is about 7.4 million. The capital city, Palembang, hosts about 1.5 million of the province's inhabitants. Jambi's population is about 3 million whereas North Sumatra's is about 12 million, according to the 2010 census. Population density on the island as a whole is just below 100 people per km^2. In the four provinces in our study, population density is fairly evenly distributed apart from coastal North Sumatra which has a higher population density than the other areas. Malay is the main language spoken in Riau and other dialects of the same family are also the main tongue in Jambi and in South Sumatra. In the interior of North Sumatra, languages of a somewhat different family dominate (Ethnologue (2009)).[9]

The current situation in the four provinces has of course been heavily influenced by general historical developments on Sumatra. Sumatra hosted several kingdoms after its initial settlement around 500 BC. One of the most dominant polities was the Buddhist Kingdom of Srivijaya, based in South Sumatra's capital Palembang. This maritime power flourished between 850-1025 AD and was a very important trading hub between east and west. After Srivijaya's decline, most of the population on Sumatra converted to Islam by the year 1300 through the influence of Arab and Indian traders. Aceh became the dominant political unit in the 16th century and resisted the increasing Dutch influence until the Aceh War 1873-1903. Trade was always a central part of the Sumatran economy, in particular during the Dutch East Indies-era when Dutch traders dominated the spice trade. In 1945, Sumatra became part of newly independent Indonesia (Ricklefs (2008)).

Although Sumatra is the main focus of our analysis, we also investigate the impact of resource windfalls among the provinces of Kalimantan, the other major oil-producing region

[9] These language are Batak Mandailing, Batak Angkola, and Batak Toba.

in the country. Kalimantan is made up of the Indonesian parts of Borneo. Before the colonial period, the southern parts of Kalimantan belonged to the Banjar sultanate (1526-1860). The Dutch colonial power increased its presence in the 19th century from their bases on Java but the current Indonesian borders of the Dutch colony were not established until in the early 20th century. Kalimantan was always considered a peripheral part of the colony. Like Sumatra, Kalimantan became part of independent Indonesia in 1945.

Kalimantan province split into three provinces in 1956; West, South, and East. The following year, South Kalimantan split into South and the geographically larger Central Kalimantan in order to give the indigenous Dayak population of Central province greater autonomy from the Muslim populations in South Kalimantan. Kalimantan hosts numerous ethnic groups of which the most important language families are Malayic, Barito, Dayak, and North Borneo.[10] A simplified description, Dayak groups dominate the interior whereas Muslim groups control the lands closer to the coast.

In terms of natural resources, Kalimantan is perhaps the richest region in the country, whereas in terms of general wealth, it is relatively undeveloped (like most of the areas outside Java). In terms of population density and the geographic dispersion of economic activity, Kalimantan is similar to the many African countries currently experiencing a resource boom. Total population in 2010 is estimated to be just below 14 million and population density is only 25 people per km^2, which can be compared with Sumatra's 100 people per km^2 and Java's equivalent Figure of over 1000 per sq. km.

East Kalimantan is the only province where oil is produced whereas no oil is produced in South, Central, and West Kalimantan. Our analysis focuses on the border area between East Kalimantan and the other three provinces, as shown in Figure 4. The widest area that we consider includes villages as far as 100 kilometers from the border. We choose this threshold because the treatment area beyond this point is mainly covered by a producing district which we have excluded from the analysis. We also study outcomes at closer distances from the border.

Figure 4 shows the geographic features of the border area. As is evident from the map, the

[10] Data is from Ethnologue (2009). Kalimantan as a whole has 74 distinct languages.

terrain is not obviously different on either side of the border. On the contrary, the topography is typically quite similar on both sides. In terms of ethnicity, our investigations show that at least the southern and central parts of the border cut right through the traditional lands of ethnic groups within the Barito language family.[11] Although we have not found specific reasons which explain why the province border is shaped as it is, we have found no information suggesting that the border was shaped by major discontinuities in colonial or pre-colonial history.

4 Data and identification strategy

4.1 Data

In this paper we make use of village data and district data. The village data come from various waves of the Indonesian Village Census (PODES), collected by the Indonesian National Institute of Statistics (BPS) every three years. We make use of the 1996, 2000, 2003 and 2006 waves.[12] The village censuses include detailed information on geographic characteristics, dwelling and wealth characteristics for the majority of the households, access to infrastructures, economic activities. The main advantage of using these data is that they cover the entire universe of Indonesian villages. This allows us to avoid problems of sample size in our study area. The second advantage is that we can merge these data with detailed information on the location of these villages.[13]

The second type of data that we use is the budget data collected by the Ministry of Finance. The data include revenue and expenditure data. The revenue components include the data on natural resource related transfers that constitute our explanatory variable of

[11] See maps on Indonesia in Ethnologue (2009).

[12] The village data are collected in preparation of larger household surveys (or censuses). Hence, the year of the PODES does not always correspond to the effective collection period. For example, the PODES 2000 data were collected during the fall 1999, the PODES 2003 during the fall 2002, while the PODES 2006 were collected during the late spring 2005.

[13] Village coordinates are available only at a specific point in time. Merging village coordinates with the village censuses is challenging because the villages have no common identifier across the different waves. Therefore we decide to track villages across waves using their name, the name of the sub-districts and districts in which they are located and detailed documents about how districts, sub-districts and villages split and aggregated over time. We successfully track about 62 percent of the villages in our baseline (1996) data.

interest.

4.2 Identification strategy

The legislative change generates automatically two groups: districts located in producing provinces and districts located in non-producing provinces. Districts belonging to the first group should have experienced a remarkable increase in their revenue driven by the oil and gas transfers. An obvious identification strategy would be to compare the two groups over time, thus applying a Difference-in-Difference (DD) strategy. Since we have the precise geo-referenced location of all the villages in the sample, we push the identification strategy one step further and adopt a spatial Regression Discontinuity (RD) design in which we compare villages facing each other from the opposite sides of the province borders.

In order to unfold the research design in a clear way, we present the identification strategy, the econometric specification and the results for the Sumatra study area. The details and the results for the Kalimantan study area are summarized in Section 6.

We consider a "large" sample including all villages within 300 kilometers from the closest border and a restricted sample including all villages within 200 kilometers from the same borders. The "large" sample includes 5107 villages (2308 treatment villages in 12 districts, 2799 control villages 14 districts), while the "restricted" sample includes 4109 villages (1949 treatment villages in 10 districts, 2160 control villages in 11 districts). Table 2A shows that, before the legislative change, treatment and control villages were broadly similar in terms of geographic, dwellings and infrastructure characteristics even when the sample includes villages relatively far from the border.[14]

Figure 5 shows the district revenue per capita over time.[15] Consistent with the implementation of Law 22/1999, districts experience a sharp increase in revenue in 2001. Figures 6 and 7 show the pattern of district revenue disaggregated by treatment/control group and

[14] Table 2B shows the comparison of treatment and control villages after we excluded villages in producing districts. The few relevant differences seem to be in terms of quality of the main road, village area and prevalence of primary schools.

[15] Since any homogeneous distance from the border cuts through several districts we further weigh the revenue by the number of villages included in the sample. District population and number of villages are obtained from the 1996 PODES data. We describe these data more in detail in the next section.

by source of revenue. The Figures support our research design: treatment areas experience a greater increase in revenue than control areas; this additional increase is driven by transfers related to natural resources.[16] Next, we look at district expenditure. Figure 8 shows that treatment areas increase their expenditure as soon as their revenue increases. This is a similar pattern to what was found in Brazil. In addition, we document strong increase in administrative expenditure (Figure 9), followed by a strong increase in transport and public work expenditure and education (Figure 10).[17]

In order to understand whether the quantity and quality of local public goods actually improve, we need to combine this data with village censuses.

4.3 Econometric specification and falsification experiments

One of the biggest advantages of our dataset is the time dimension. First, it allows to compare villages close to the border before and after the legislative change. Second, joint with our work on tracking villages over time, it allows us to include village fixed effects in the analysis. Third, joint with the availability of two waves of data before the legislative change takes place, it allows us to estimate the impact of the legislative change on treatment villages before the legislative change actually takes place (falsification experiment).

While a typical spatial RD design requires all determinants of the outcome of interest to change smoothly at the boundary (so that villages just outside the boundary are an appropriate control group to villages just inside it), the availability of village characteristics before and after the legislative changes requires us to assume that only *time-varying* determinants of the outcome of interest change smoothly at the boundary.[18] This assumption is significantly

[16] Treatment and control groups still exhibit great differen even if we exclude the producing districts (not reported).

[17] It is not trivial to follow sector expenditure over time because the Ministry of Finance changed the budget structure in 2003. This creates two problems. First, not all districts switched to the new system at the same time. Second, the expenditure categories with the new system do not match well the old categories (i.e., it is very difficult to reconstruct the entire time series using only one reporting system). In our case before 2003 routine expenditure for sectors like education (e..g teachers' wages) fell into the administrative category, while building new schools fell into the education section. Along the same lines, there is a relationship between transportation expenditure in 2001-2002 and public work expenditure 2003-2005, although the details are unclear.

[18] To our knowledge, the only other empirical application combining a spatial RD design with a time dimension is Lemieux and Milligan (2008).

weaker if one was concerned with persistent differences between the two groups.

With the sample of villages so restricted, we estimate the impact of the legislative change on treatment villages using the following econometric specification:

$$c_{idt} = \alpha_{1,i} + \gamma_{1,1999}(T_d * 1999) + \sum_{t\in\{2002,2005\}} \left[\gamma_{1,t}\left(T_d * d_t\right)\right] + d_t + \varepsilon_{idt} \quad (1)$$

where c_{idt} is the outcome in village i, district d at time t, α_i is a village fixed-effect, T_d is a measure of the resource windfall, d_t is the year fixed effect, and ε_{idt} is the error term clustered at the district level.

We will use two different measures of the resource windfall: a simple binary variable indicating whether the district is located within a producing province, and a continuous variable capturing the average per capita oil and gas transfers that the district government received during the current and the two previous periods.[19] Our outcomes of interest measure the amount of public goods that villagers have access to: for education we use binary variables indicating whether the village has a primary school, whether it has a junior-high school and whether it has a senior-high school; for health we use binary variables indicating whether the village has a maternity hospital/house, whether it has a health center and whether the majority of the households have access to piped water; for transportation we use a binary variable indicating whether the majority of the traffic is through land (as opposed to water), whether the main village road is paved and whether the village has a bus terminal; other infrastructures are whether the village has a public phone, whether it has a post office and whether it has a permanent market. The key parameter of interest are $\gamma_{1,2002}$ and $\gamma_{1,2005}$. These two coefficients capture the impact of the resource windfall on public good provision as long as there are no differences between treatment and control villages (other than the resource windfall) that vary over time and are correlated with public good provision (identification assumption). If $\gamma_{1,2005} > 0$ or even $\gamma_{1,2002} > 0$, this means that treatment villages experienced

[19] Including the two transfers preceding the current one in our measure seems appropriate not only because the effect of the resource windfall may take place with one or more lags, but also because districts received not transfers in 2000. Hence, the measure incorporates by construction the fact that the impact in 2002 may be weaker than the impact in 2005.

a greater increase in public goods than the control group after the legislative change. The other parameter of interest is $\gamma_{1,1999}$, which works as a falsification experiment. If $\gamma_{1,1999} = 0$, then treatment villages experienced no variation in public goods relative to the control group before the legislative change. This would be consistent with the identification assumption being valid. On the contrary, a coefficient estimate different from zero would shed some doubt on the validity of the identification assumption.

Specification (1) essentially corresponds to a Difference-in-Difference (DD) specification. Since we have detailed information on the geographic location of each village in the dataset, we can specify further our econometric model using a spatial Regression Discontinuity (RD) design. Like Dell (2010) we have not enough units close to the border to specify a fully flexible local linear regression. Hence, we turn to the following semi-parametric specification:

$$c_{idbt} = \alpha_{2,ib} + \gamma_{2,1999}(T_d * 1999) + \sum_{t \in \{2002,2005\}} \left[\gamma_{2,t}\left(T_d * d_t\right)\right] + [f(location_i) * d_t] + (\Phi_b * d_t) + \varepsilon_{ibdt} \tag{2}$$

where $f(location_i)$ is a function of the geographic location of the village, Φ_b is binary indicator for the boundary and ε_{ibdt} clustered at the district level. Since econometric theory (and practice) does not provide precise indications on which functional form is superior in a spatial RD design, we use three different specifications: a cubic polynomial of the latitude and the longitude of the village; the distance of the village to the closest border; a cubic polynomial of the distance of the village from the border.[20] The interaction between the segment indicators and the year indicators ($\Phi_b * d_t$) imply that we control for segment-year fixed effects, rather than just for year fixed effects. Controlling for segment-year fixed effects means that the comparison between villages on different sides of the boundaries that identifies our coefficient estimates of interest is restricted to those villages "facing each other", i.e., lying on different sides of the same boundary. This could be important because our study area includes three different boundaries that are located far away from each other. The

[20] In the results we report only the coefficient estimates associated with the cubic polynomial of the distance to the border for the sake of brevity. The results for the other two alternatives are typically very similar and are available upon request.

two coefficients of interest ($\gamma_{2,2002}$ and $\gamma_{2,2005}$) capture the impact of the resource windfall on public good provision as long as there are no differences between treatment and control villages (other than the resource windfall) that vary over time, are correlated with public good provision and vary discontinuously across the border. The identification assumption is weaker than in specification (1) because the inclusion of segment-year fixed effects and the flexible function of village location allows possible differences in local markets not to confound our effect of interest.[21]

Finally, since our sample of districts includes both producing and non-producing districts, and local governments may behave very differently across the two categories, we will also estimate specification (2) controlling for a binary indicator for producing districts interacted with a full set of time dummies. This will ensure that our coefficients of interest (γ_{2002} and γ_{2005}) capture the impact of the resource windfall on local government behavior rather than the direct impact of oil and gas extraction on the local economy.

5 Results

Table 3A shows the impact of the resource windfall on transportation infrastructures, i.e., whether the majority of the traffic runs through land (Columns 1-4), whether the village road is paved (Columns 5-8) and whether the village has a bus terminal (Columns 9-12). Panel A shows the results associated with the binary treatment indicator. The effect of the revenue windfall on the likelihood that most of the traffic runs through land is close to zero in all specifications. The effect of the revenue windfall on road quality seems to be negative before decentralization, positive and relative large immediately afterwards, and positive and small in the medium term.[22] None of the coefficient estimates is consistently significant. Panel B and

[21] Controlling for distance to the border is similar to controlling for distance to the extraction points. We expect oil extraction to influence the markets closeby. However, we also expect such influence to fade away smoothly with distance from the extraction points (Aragon and Rud (Forthcoming)). Treatment villages are, on average, closer to these extraction points than control villages, but any possible direct (time-varying) impact of this difference on public good provision should be captured by the distance to the border.

[22] Specifying the function of geographic location as (linear) distance to the border or a cubic polynomial of latitude and longitude does not affect the results. This holds true for all other results in this section unless otherwise specified.

C show the results associated with per capita oil and gas transfers averaged over the current and the previous two years. The effect of the revenue windfall on road quality is positive and significant throughout all specifications. The magnitude is rather small: an increase in per capita oil and gas revenue of one standard deviation (15 USD) increases the probability of having the road paved by 3 percentage points, i.e., about 5 percent of the pre-decentralization average. In addition, the resource windfall seems to have no effect on the likelihood of having a bus terminal (Columns 5-8).

Table 3B shows the impact of the resource windfall on education infrastructures, i.e., whether the village has a primary school (Columns 1-4), a junior-high school (Columns 5-8) or a senior-high school (Columns 9-12). The revenue windfall seems to have no effect on the likelihood of having a primary school. On the other hand, it seems to have a consistent positive impact on the likelihood of having a junior-high school: being located in a producing province is associated with an increase of 8.3 percentage points of having a junior-high school in 2002 and 9.9 points in 2005, i.e., an increase of (respectively) 26 and 31 percentage points relative to the pre-decentralization average. The coefficient estimates associated with the continuous measure of the revenue windfall (Panel B and C) confirm the direction and significance of the impact. An increase in per capita oil and gas revenues of one standard deviation is associated with an increase of 3 percentage points, i.e., an increase of 9.3 points relative to the pre-decentralization average. On the contrary, the results for senior-high schools are mixed: the coefficient estimates are positive but relatively large only in the second period and not always significant. An increase in per capita oil and gas revenues of one standard deviation is associated with an increase of 1.5 percentage points, i.e. an increase of 8.3 points relative to the pre-decentralization average. Overall, the resource windfall has the strongest impact on junior-high schools. This is consistent with the what we know about education in Indonesia: primary education is almost universal (87 percent of villages have a primary school); junior-high schools are widespread but not nearly as much as primary schools (junior-high schools are present in 32 percent of the villages) notwithstanding the increase in mandatory education to the first nine grades adopted since the early 1990s; senior-high schools are relatively rare (they are present in 16 percent of the villages).

Table 3C shows the impact of the resource windfall on health infrastructures, i.e., whether the village has a maternity house (Columns 1-4), a health center (Columns 5-8), or whether the majority of the households has access to piped water (Columns 9-12). We find no evidence of an increase in health infrastructures following the increase in oil and gas revenues.

Table 3D shows the impact of the resource windfall on whether the village has a public phone (Columns 1-4), a post office (Columns 5-8) or a permanent market (Columns 9-12). The resource windfall does not seem to have led to more communication infrastructures, although the results for the post office are mixed. On the contrary, it seems to have led to better trade infrastructures: being located in a treatment village is associated with an increase of 7.7 percentage points on the likelihood of having a permanent market in 2002 and 7.9 in 2005, i.e., 55 and 57 percent relative to the pre-decentralization average. The coefficient estimates associated with the continuous measure of the revenue windfall (Panel B and C) confirm the direction and significance of the impact. An increase in per capita oil and gas revenues of one standard deviation is associated with an increase of 3 and 1.5 percentage points, i.e., an increase of, respectively, 20 and 10 percentage points relative to the pre-decentralization average.

In order to make sure that the producing districts are not driving the evidence found so far, we re-estimate the previous models controlling for an interaction between a producing indicator and the time dummies. By controlling for the producing districts, our coefficient estimates should capture uniquely the effect of the revenue windfall on local government behavior without any obvious direct impact of oil extraction. Table 4 shows the results for selected outcomes: whether the village has a paved road (Columns 1-4), whether it has a junior-high school (Columns 5-8), whether it has a permanent market (Columns 9-12). The coefficient estimates essentially confirm the previous results: the revenue windfall led to more junior-high schools and more permanent markets, while the evidence for road quality is, again, mixed.

6 Resource windfall and public goods in Kalimantan

As discussed in Section 3, there is only one other area in Indonesia where the oil and gas extraction takes place at a scale so high that it leads to relevant oil and gas transfers to districts located within the same province of the producing district. This area is Kalimantan. Oil and gas extraction takes place almost exclusively in East Kalimantan. Hence, our analysis focuses on the border area between East Kalimantan and the other three provinces (Figure 4). The study area is not as large as for Sumatra. Hence, we expect to have a lower power to detect any change in public good provision following the redistribution of resource revenues. On the other hand, replicating the analysis in a different region of Indonesia may yield interesting insights, for example, in terms of external validity, since Kalimantan and Sumatra differ in many aspects.

The widest area that we consider includes villages as far as 100 kilometers from the border. We choose this threshold because the treatment area beyond this point is mainly covered by a producing district which we want to exclude from the analysis. Along the lines of the analysis for Sumatra, we gradually restrict the study area to villages within 75, 50 and 25 kilometers from the border. Our final sample includes: 1,551 villages for the 100 kilometer sample (275 treatment, 1,276 control); 1,174 villages for the 75 kilometer sample (222 treatment, 952 control); 589 villages for the 50 kilometer sample (158 treatment, 431 control); 187 villages for the 25 kilometer sample (52 treatment, 135 control).[23] Table 5 shows the comparison of treatment and control villages in terms of a wide range of geographic and demographic characteristics, as well as in terms of dwellings and public infrastructures. Treatment and control villages show some differences, but they tend to fade away as we get closer to the border.

Figure 11-16 shows the pattern of revenue, revenue components and expenditure across treatment and control villages over time. The pattern confirms that treatment villages did experience a resource windfall following the implementation of the fiscal decentralization law even stronger than in the Sumatra study area.

[23] For Kalimantan we manage to track over 90 percent of the villages over time. Hence, we are much more confident in the quality of the data throughout the entire analysis than we did for Sumatra.

We turn to the econometric analysis. It is important to keep in mind that a recent econometrics literature (see Cameron and Miller (2011) and references therein) has found that clustering the standard errors with less than 30 clusters can lead to underestimating the true standard errors. The problem can be particularly severe with less than 10 clusters and generally fades away as the number of clusters increases. Based on Monte-Carlo simulations in Cameron, Gelbach, and Miller (2008), the econometric analysis that we carried out for the Sumatra study area should be still valid (since we used 26 and 23 clusters). However, the analysis for the Kalimantan study area will be based on a number of clusters ranging from 15 in the largest sample to 8 in the smallest.

We re-estimate specification (1) and (2) for this study area. Due to data constraints, we estimate only the specifications using the binary treatment measure. Table 6A, Panel A, shows the results for transportation infrastructures: whether the majority of the traffic runs on land (Columns 1-5); whether the main road is paved (Columns 6-10); whether the village has a bus terminal (Columns 11-15). The coefficient estimates of interest are positive, large and significant for whether the majority of the traffic is through land (rather than water): being located in a treatment village is associated with an increase of 15 percentage points in 2002 and 25 points in 2005 of having the majority of the traffic through land, i.e., an increase of 23 and 38 percentage points relative to the pre-decentralization average (about 66 percent). On the other hand, the resource windfall does not seem associated with an increase in quality of the road (whether it is paved or not) and public transport facilities (whether there is a bus terminal in the village).

Panel B shows the results associated with education infrastructures. As for Sumatra, the resource windfall does not seem to be associated with an increase in primary schools (Columns 1-5), while it does seem to be associated with an increase in junior-high schools (Columns 6-10) and senior-high schools (Columns 11-15). The increase in likelihood of having a junior-high school in the village is 2.8 percentage points in 2002 and 9.3 in 2005, i.e., an increase of 18 and 60 percent relative to the pre-decentralization average (15.6 percent). The increase in likelihood of having a senior-high school in the village is 2.2 percentage points in 2002 and 10.6

in 2005, i.e., an increase of 56 and 270 percentage points relative to the pre-decentralization average (3.9 percent).

Table 6B, Panel A, shows the results for health infrastructures. The resource windfall seems to be associated with an increase of 2.8 percentage points in 2002 and 2.6 in 2005, i.e., 700 and 650 percentage points relative to the pre-decentralization average (0.4 percent). However, notice that the pre-decentralization effect is almost half as large as the coefficient estimates of interest, so one must use caution to interpret these estimates as evidence of such a strong effect of the resource windfall of maternity hospitals. Indeed, having a maternity hospital in the village seems to be an event so rare that few observations may be driving the entire result. This is not the case for the second health outcomes: whether the village has a health center. There appear to be few health centers in the study area, but having one is not such a rare event as it was for the maternity hospital (the pre-decentralization average is 8.1 percent). The resource windfall is associated with no effect in 2002 and an increase of 8.6 percentage points in 2005, i.e., 106 percent relative to the pre-decentralization average. There does not seem to be any effect on access to piped water (the coefficient estimate is positive and significant in 2002, but it is not robust across the various specifications).

Panel B shows the results for other infrastructures. The resource windfall is associated with an increase in whether the village has a public phone (Columns 1-5). However, as for the maternity hospital, the coefficient estimates appear unreasonably large and this may be due to the scarcity of this facility in Kalimantan (the pre-decentralization average is 2.3 percent). On the other hand, the resource windfall seems to have increase the presence of post offices (temporary or permanent): the effect is about zero in 2002 but there is an increase of 2.6 in 2005, i.e., an increase of 40 percent relative to the pre-decentralization average (6.5 percent). Finally, the resource windfall does not seem to have a clear effect of trade facilities: the coefficient estimates associated with whether there is a permanent market in the village vary in magnitude and significance across specifications.

Table 7 shows the results for some selected outcomes once we control for the (few) villages located in producing districts. The coefficient estimates are virtually identical to those previously found.

7 Conclusions

In this paper, we study the impact of a fiscal decentralization programme in Indonesia that provided producing provinces with a greater share of resource revenue from oil and gas. Our main research question is whether this change actually led to an increase in the provision of local public goods like health and education. The previous literature on natural resources and economic development suggests several reasons why resource windfalls might actually not contribute to an improved supply of public goods.

In order to make our study as similar to a natural experiment as possible, we restrict our analysis to natural-resource rich regions Sumatra and Kalimantan and to comparing villages close to the border between oil producing and non-producing provinces. Our empirical analysis employs a regression discontinuity design where we use different distances to border as a restriction for inclusion in our treatment and control groups. Our results suggest that high school facilities tended to improve in the treated villages, in particular 3-4 years after decentralization. This finding seems robust across the two study areas. Otherwise the revenue windfall is associated with an increase in trade infrastructures in Sumatra (but not in Kalimantan), while it is associated with an increase in road, health, and communication infrastructures in Kalimantan (but not in Sumatra). We found no evidence of a decrease in public goods. Hence, we find no indications of a curse of natural resources but rather of beneficial or no effects of the resource windfall.

We believe the Indonesian fiscal decentralization program might provide poor, resource-abundant countries with an interesting policy experiment that has not previously been widely tested throughout the world. There are however numerous issues that remain to be studied within our Indonesian context. For instance, it would be useful to reach a stronger understanding of the political economy at local level. Why do an increase in resource rents lead to an increase in certain public goods and not in others? In future work, we also hope to obtain better measures of public good *quality*. For instance, it would be interesting to analyze whether the improved facilities for high schools are also associated with an improved pupil attendance or stronger test scores. This is left for future work.

References

Akai, N., and M. Sakata (2002): "Fiscal decentralization contributes to economic growth: evidence from state-level cross-section data for the United States," *Journal of Urban Economics*, 52(1), 93–108.

Aragon, F. M., and J. P. Rud (Forthcoming): "Natural Resources and Local Communities: Evidence from a Peruvian Gold Mine," *American Economic Journal: Economic Policy.*

Brollo, F., T. Nannicini, R. Perotti, and G. Tabellini (Forthcoming): "The Political Resource Curse," *American Economic Review.*

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 90(3), 414–427.

Cameron, C. A., and D. L. Miller (2011): "Robust Inference with Clustered Data," in *Handbook of Empirical Economics and Finance*, ed. by A. U. Giles, and D. E., pp. pp.1–28. CRC Press.

Caselli, F., and G. Michaels (Forthcoming): "Do Oil Windfalls Improve Living Standards? Evidence from Brazil," *American Economic Journal: Applied Economics.*

Collier, P. (2008): *The Bottom Billion: Why the Poorest Countries are Failing and What Can Be Done About It.* Oxford University Press, New York.

Dalgaard, C.-J., and O. Olsson (2008): "Windfall Gains, Political Economy and Economic Development," *Journal of African Economies*, 17(suppl 1), 72–109.

Davoodi, H., and H.-f. Zou (1998): "Fiscal Decentralization and Economic Growth: A Cross-Country Study," *Journal of Urban Economics*, 43(2), 244–257.

DELL, M. (2010): "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 78(6), 1863–1903.

DUFLO, E., R. GLENNERSTER, AND M. KREMER (2007): "Chapter 61 Using Randomization in Development Economics Research: A Toolkit," in *Handbook of Development Economics*, ed. by T. P. Schultz, and A. S. John, vol. Volume 4, pp. 3895–3962. Elsevier.

ETHNOLOGUE (2009): "Ethnologue: Languages of the World," Lewis, M. Paul.

GYLFASON, T. (2001): "Natural resources, education, and economic development," *European Economic Review*, 45(4âĂŞ6), 847–859.

LEITE, C. A., AND J. WEIDMANN (1999): "Does Mother Nature Corrupt? Natural Resources, Corruption, and Economic Growth," .

LEMIEUX, T., AND K. MILLIGAN (2008): "Incentive effects of social assistance: A regression discontinuity approach," *Journal of Econometrics*, 142(2), 807–828.

MEHLUM, H., K. MOENE, AND R. TORVIK (2006): "Institutions and the Resource Curse*," *The Economic Journal*, 116(508), 1–20.

MONTEIRO, J., AND C. FERRAZ (2009): "Resource Booms and Voracious Politicians: The Effects of Oil Shocks on Patronage, Rent-Seeking, and Elections," .

REINIKKA, R., AND J. SVENSSON (2004): "Local Capture: Evidence from a Central Government Transfer Program in Uganda," *The Quarterly Journal of Economics*, 119(2), 679–705.

RICKLEFS, M. C. (2008): *History of Modern Indonesia Since c. 1200: Fourth Edition*.

ROBINSON, J. A., R. TORVIK, AND T. VERDIER (2006): "Political foundations of the resource curse," *Journal of Development Economics*, 79(2), 447–468.

SACHS, J. D., AND A. M. WARNER (1997): "Natural Resource Abundance and Economic Growth," .

——— (2001): "The curse of natural resources," *European Economic Review*, 45(4âĂŞ6), 827–838.

Sala-i Martin, X., and A. Subramanian (2003): "Addressing the Natural Resource Curse: An Illustration from Nigeria," .

Skoufias, E., A. Narayan, B. Dasgupta, and K. Kaiser (2011): "Electoral accountability, fiscal decentralization and service delivery in Indonesia," .

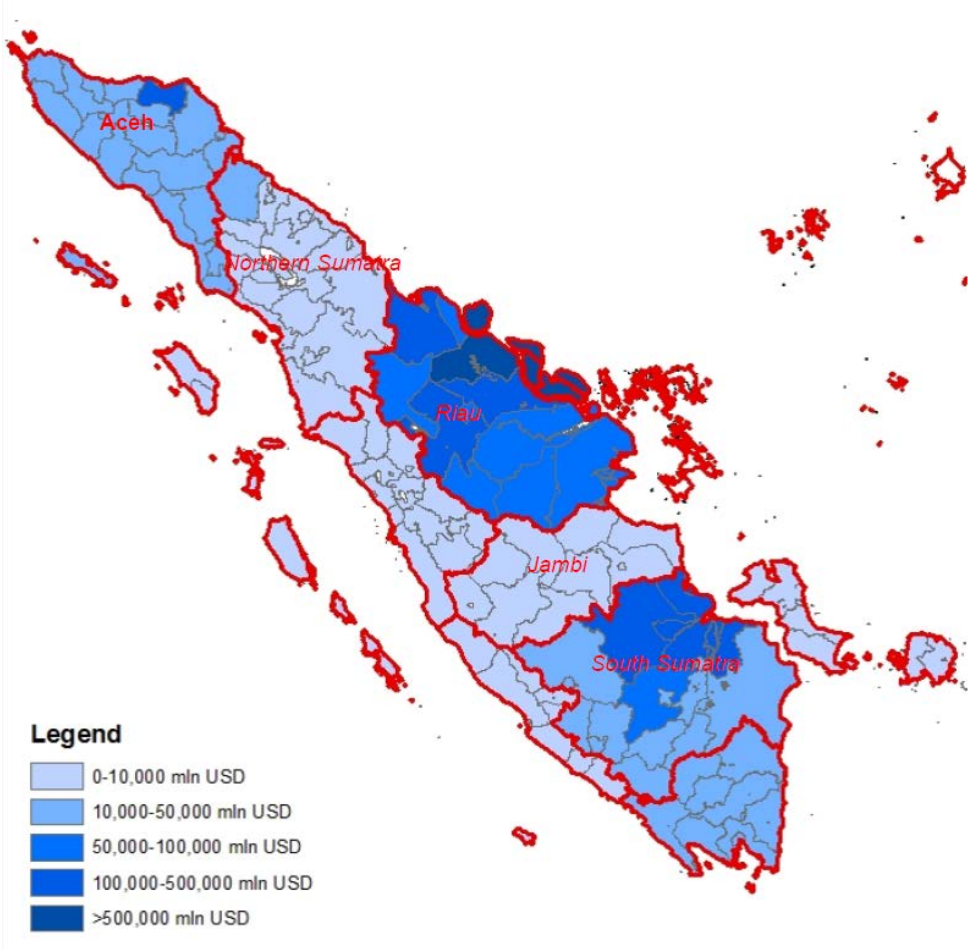
Torvik, R. (2002): "Natural resources, rent seeking and welfare," *Journal of Development Economics*, 67(2), 455–470.

van der Ploeg, F. (2011): "Natural Resources: Curse or Blessing?," *Journal of Economic Literature*, 49(2), 366–420.

Vicente, P. C. (2010): "Does oil corrupt? Evidence from a natural experiment in West Africa," *Journal of Development Economics*, 92(1), 28–38.

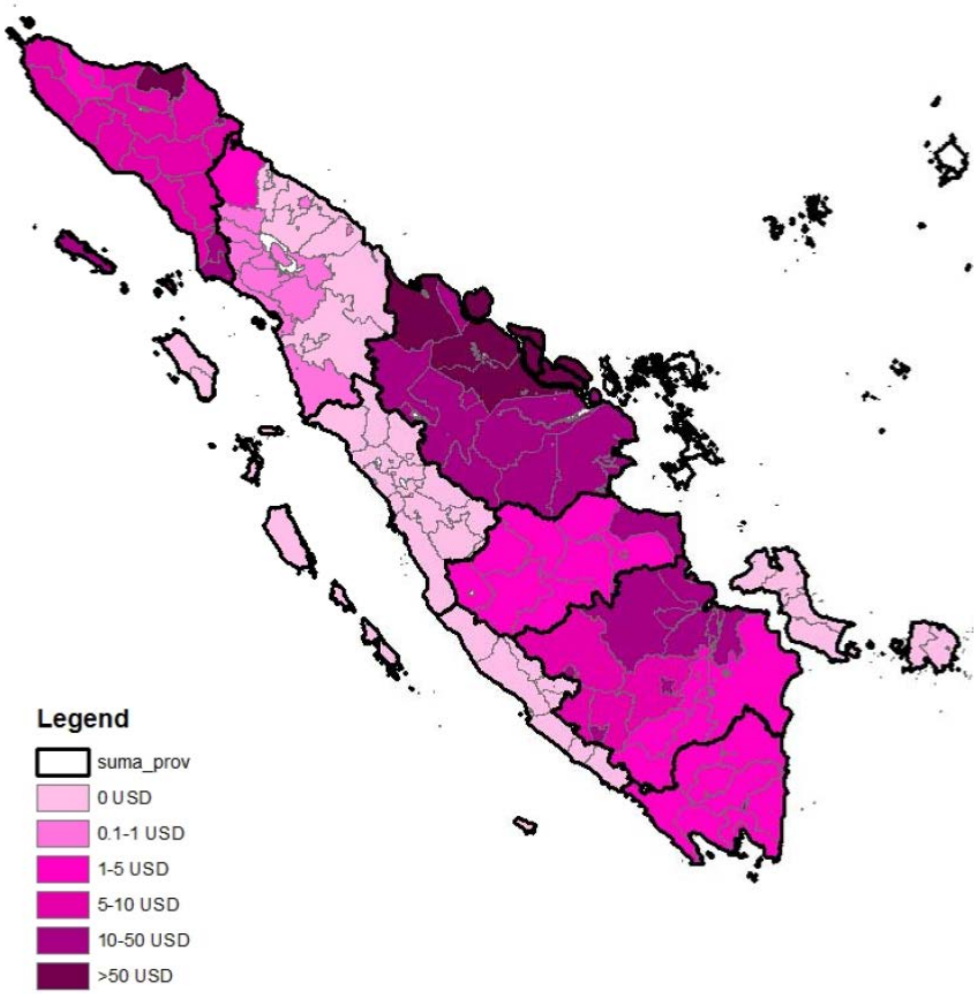
Zhang, T., and H.-f. Zou (1998): "Fiscal decentralization, public spending, and economic growth in China," *Journal of Public Economics*, 67(2), 221–240.

Figure 1: Magnitude of oil and gas transfers in absolute terms in 2002 on Sumatra



Note: The thick lines show province borders whereas the thin lines show district borders. The names of the four provinces included in the study are in italics.

Figure 2: Magnitude of per capita oil and gas transfers in 2002 on Sumatra

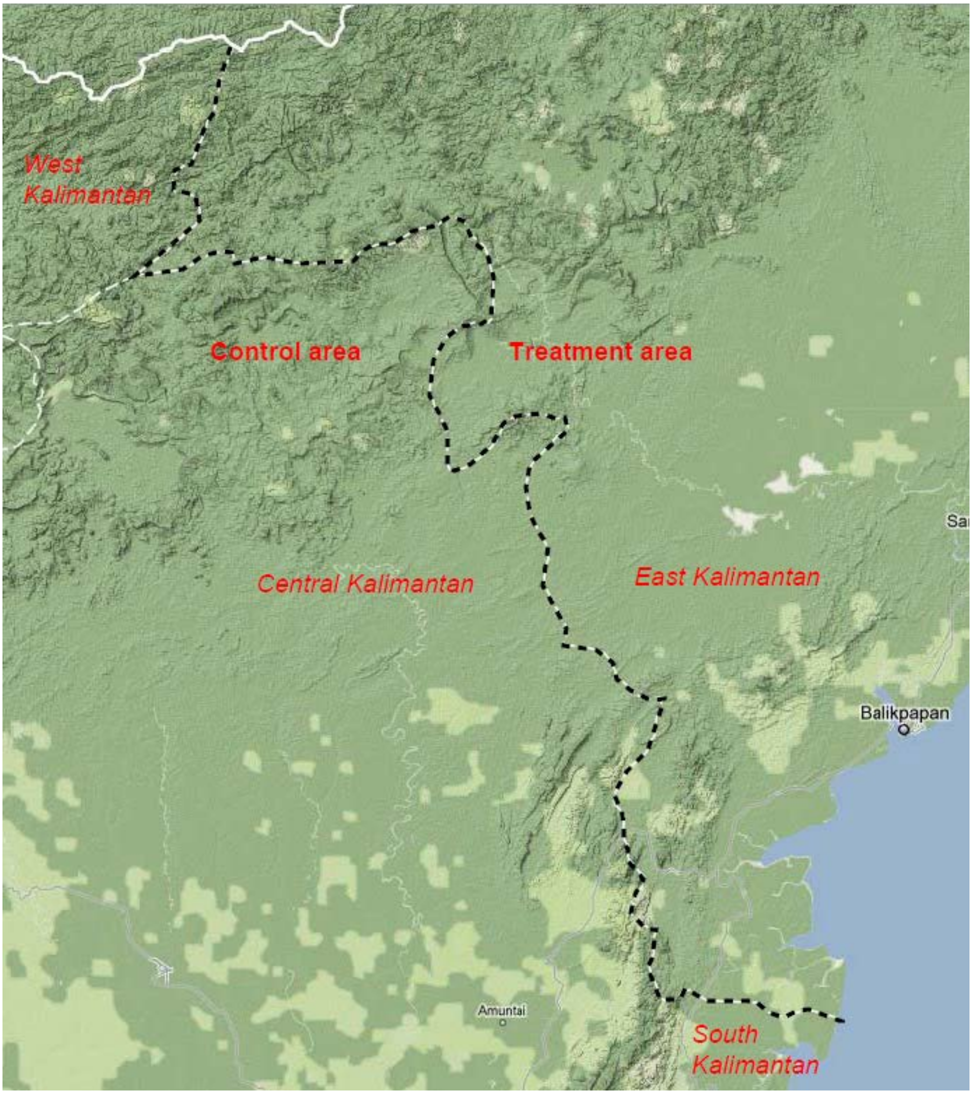


Note: The thick lines show province borders whereas the thin lines show district borders.

Figure 3: Borders between treatment (Riau, South Sumatra) and control areas (Northern Sumatra and Jambi) on Sumatra

Note: The three dotted lines show the borders exploited in the empirical study between treatment and control areas. The northeastern dotted line is between Riau (treatment) and North Sumatra (control), the central line between Riau (treatment) and Jambi (control), and the southernmost line is between South Sumatra (treatment) and Jambi (control). The black lines show the borders to provinces not included in the study.

Figure 4: Border between treatment (East Kalimantan) and control areas (West, Central, and South Kalimantan)



Note: The dotted lines show the borders exploited in the empirical study between the treatment area East Kalimantan and the control areas West, Central, and South Kalimantan.

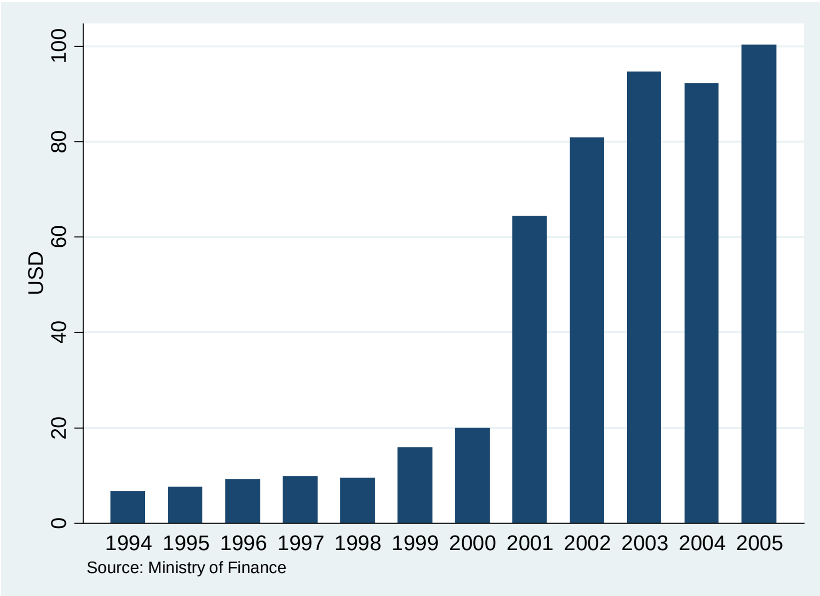


Figure 5: Evolution of per capita district revenue over time (Sumatra).

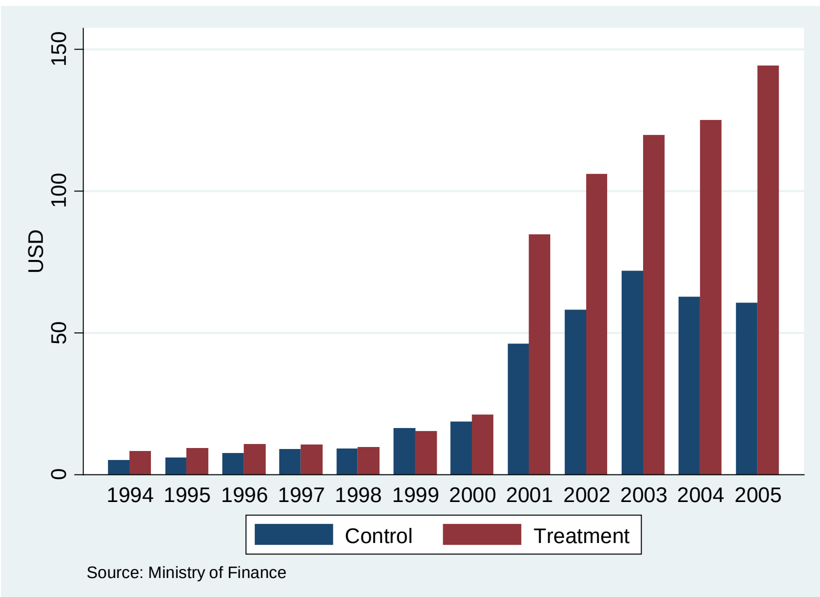


Figure 6: Per capita district revenue in treatment and control areas (Sumatra).

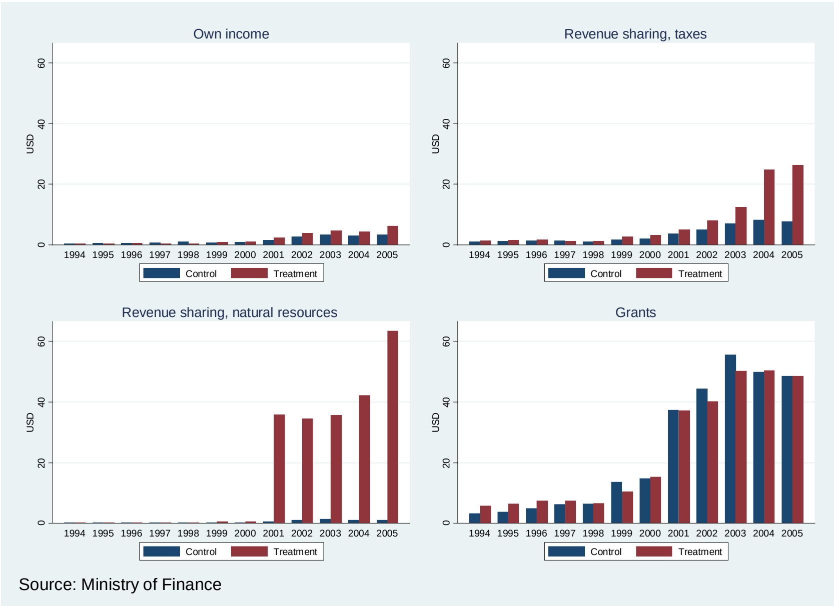


Figure 7: Per capita district revenue components in treatment and control areas (Sumatra) .

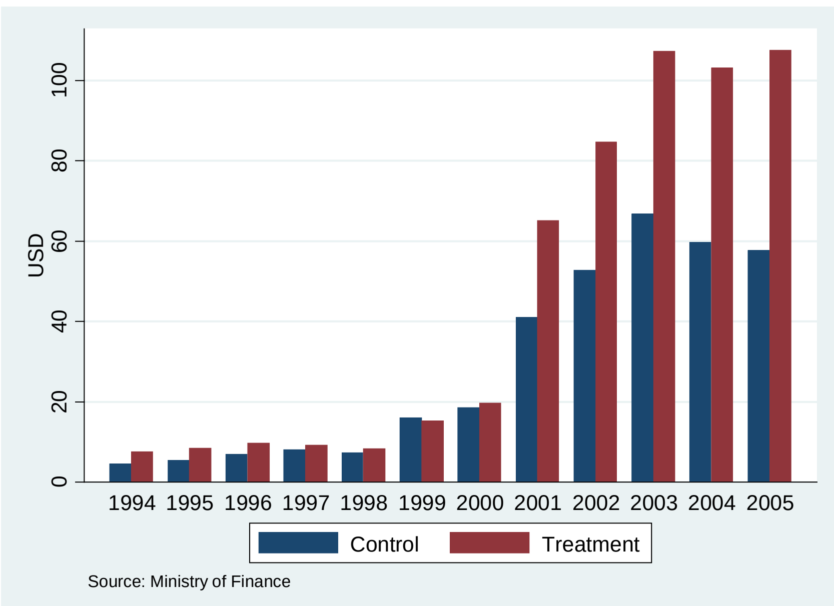


Figure 8: Per capita district expenditure in treatment and control areas (Sumatra).

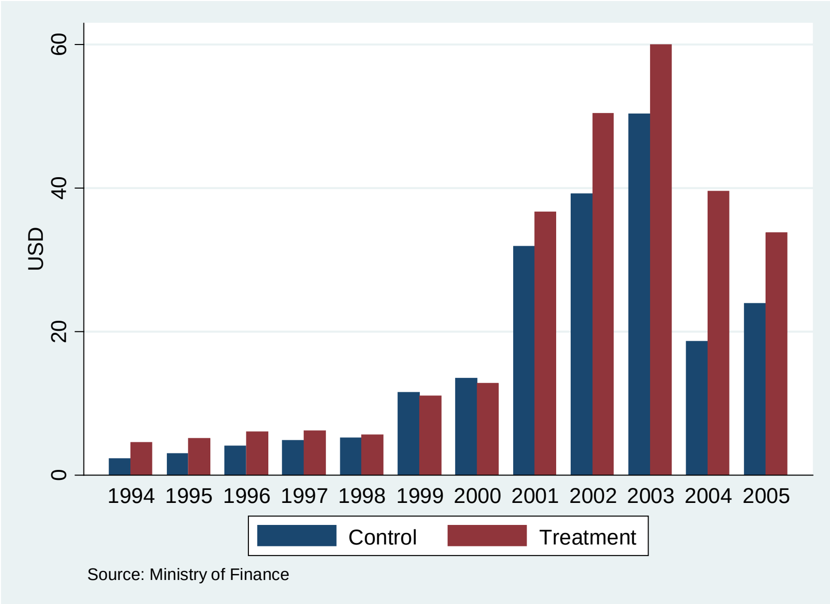


Figure 9: Per capita district administrative expenditure in treatment and control areas (Sumatra).

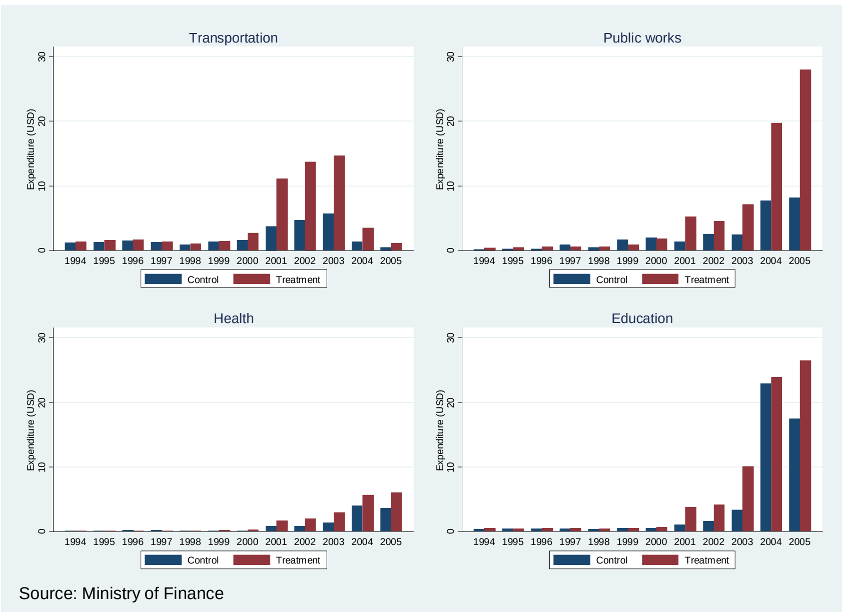


Figure 10: Per capita district expenditure, other primary components (Sumatra).

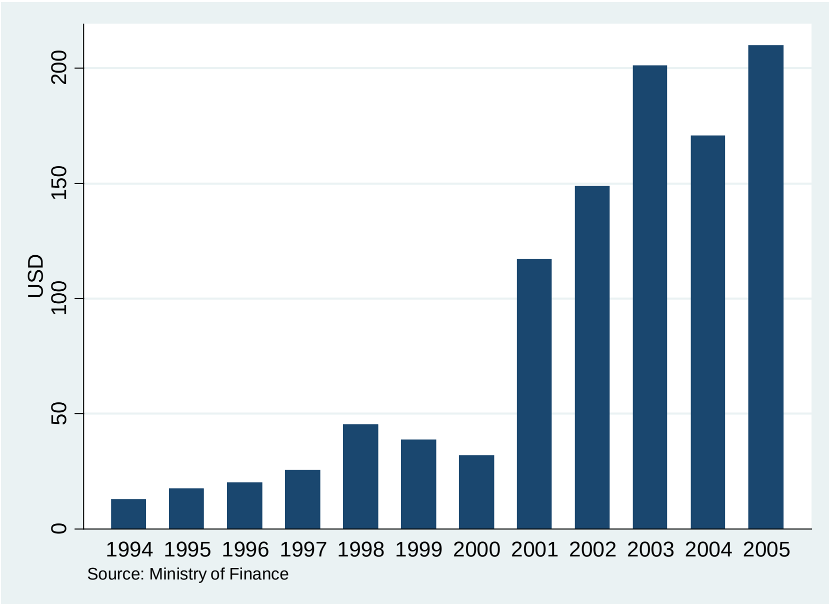


Figure 11: Evolution of per capita district revenue over time (Kalimantan).

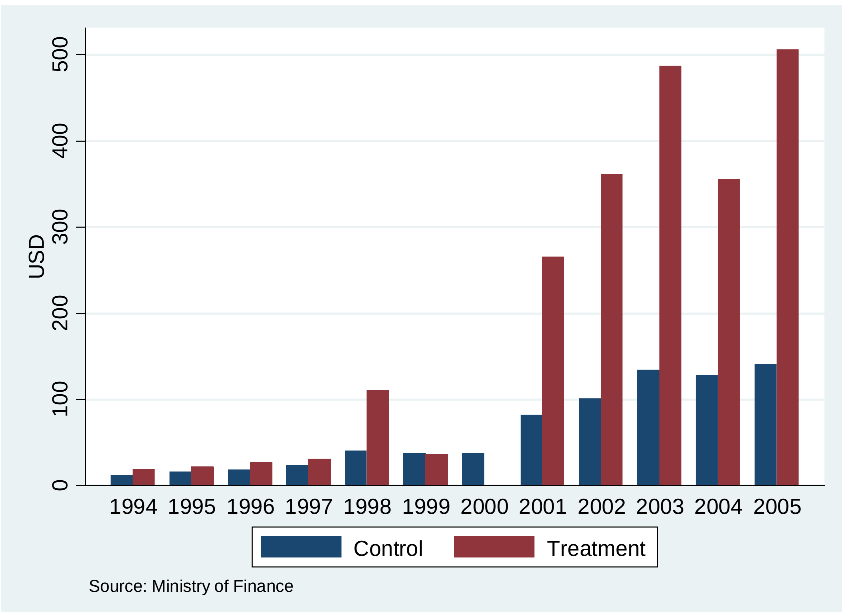


Figure 12: Per capita district revenue in treatment and control areas (Kalimantan).

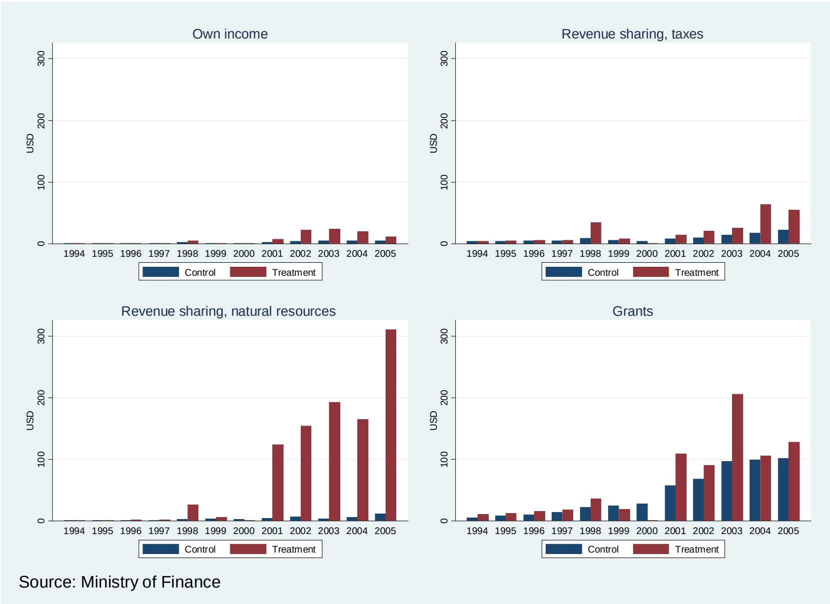


Figure 13: Per capita district revenue components in treatment and control areas (Kalimantan) .

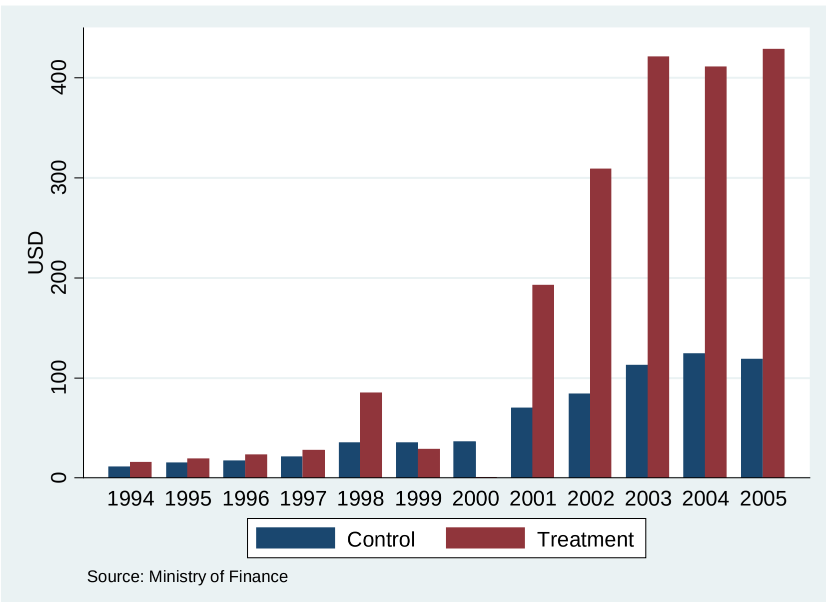


Figure 14: Per capita district expenditure in treatment and control areas (Kalimantan).

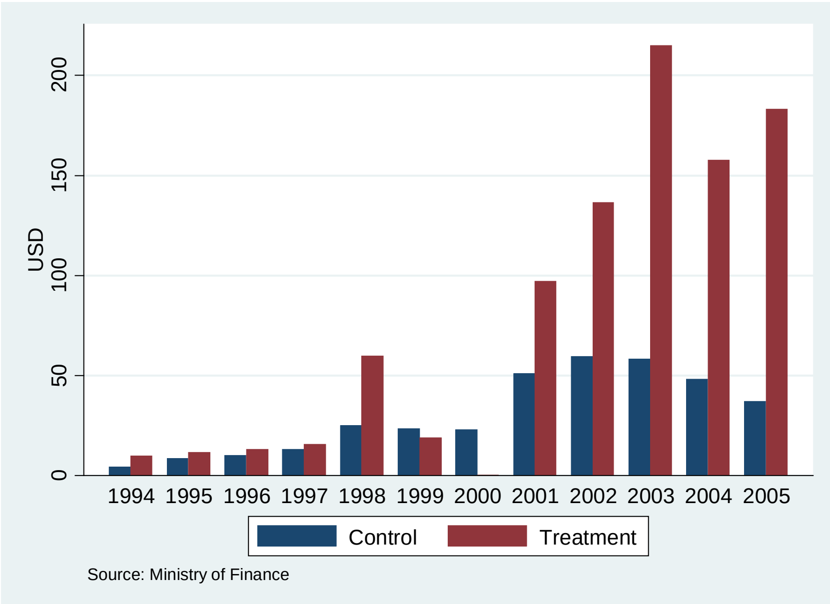


Figure 15: Per capita district administrative expenditure in treatment and control areas (Kalimantan).

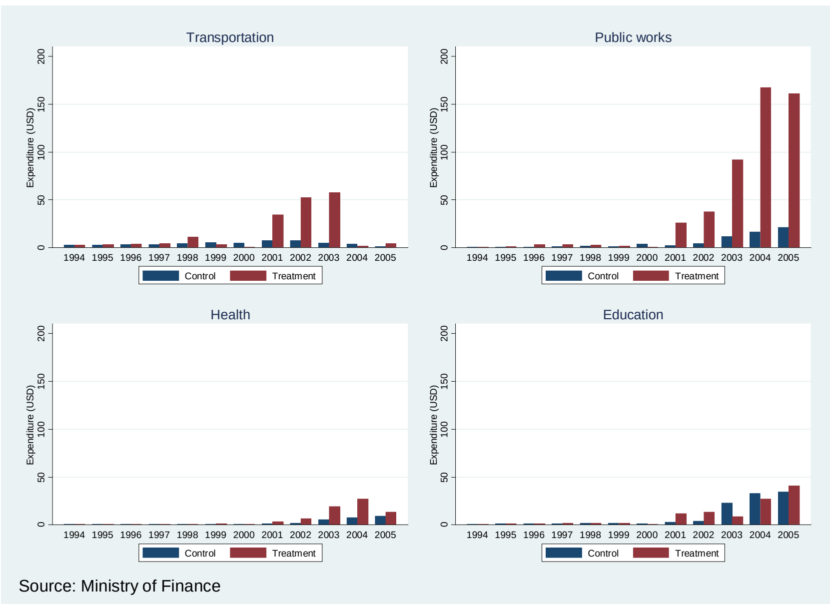


Figure 16: Per capita district expenditure, other primary components (Kalimantan).

TABLE 2A

DESCRIPTIVE STATISTICS: SUMATRA AREA

Sample	<300 kilometers			<200 kilometers		
Treatment	Oil-rich	Oil-scarce	Difference	Oil-rich	Oil-scarce	Difference
Statistic	mean	mean	t-stat	mean	mean	t-stat
Village located on the coast	0.065	0.043	(0.493)	0.062	0.035	(0.646)
Village located in a valley	0.026	0.078	(-1.414)	0.025	0.094	(-1.693) *
Village located in a hilly area	0.099	0.129	(-0.552)	0.071	0.131	(-1.075)
Village located in a plane	0.814	0.75	(0.838)	0.847	0.739	(1.288)
Urban village	0.097	0.176	(-1.003)	0.108	0.119	(-0.158)
Official urban village	0.118	0.154	(-0.459)	0.135	0.116	(0.291)
Population	3109.682	3089.259	(0.020)	2793.325	2447.605	(0.378)
Number of households	559.724	587.714	(-0.153)	572.789	449.206	(0.723)
Number of households in agriculture	303.38	267.816	(0.513)	291.142	260.004	(0.375)
Electricity in the village	0.863	0.892	(-0.611)	0.879	0.87	(0.157)
Share of households with electricity	0.331	0.533	(-2.576) ***	0.355	0.491	(-1.614)
Public phone in the village	0.092	0.135	(-0.633)	0.1	0.081	(0.335)
Number of households with phone[1]	0.019	0.022	(-0.287)	0.021	0.014	(0.638)
Main road is lighted	0.237	0.467	(-2.021) **	0.259	0.371	(-1.063)
Majority uses LPG/Kerosene for cooking	0.114	0.194	(-1.019)	0.124	0.121	(0.053)
Majority litters in bin, then delivered	0.123	0.134	(-0.166)	0.138	0.104	(0.612)
Majority households has private toilet	0.348	0.468	(-0.799)	0.34	0.392	(-0.317)
Majority households has public toilet	0.601	0.437	(1.499)	0.61	0.492	(0.995)
Streaming sewage system in village	0.588	0.679	(-1.167)	0.57	0.653	(-0.900)
Share of permanent dwellings	0.149	0.188	(-0.779)	0.148	0.154	(-0.127)
Slum in village	0.128	0.063	(1.951) *	0.138	0.051	(2.404) **
Share of households living in slum	0.02	0.007	(2.596) ***	0.022	0.006	(3.106) ***
Primary school in the village	0.915	0.815	(1.078)	0.91	0.795	(1.026)
Junior-high school in the village	0.286	0.335	(-0.591)	0.29	0.288	(0.017)
Senior-high school in the village	0.127	0.18	(-0.979)	0.128	0.137	(-0.188)
Number of primary schools in the village[1]	2.404	2.838	(-0.584)	2.416	2.655	(-0.273)
Number of junior-high schools in the village[1]	0.446	0.597	(-0.832)	0.456	0.462	(-0.036)
Number of senior-high schools in the village[1]	0.211	0.352	(-1.169)	0.21	0.248	(-0.410)
Hospital in the village	0.017	0.044	(-1.283)	0.019	0.027	(-0.561)
Maternity house in the village	0.045	0.125	(-1.457)	0.049	0.069	(-0.520)
Health center in the village	0.097	0.111	(-0.395)	0.1	0.095	(0.140)
Doctor in the village	0.145	0.177	(-0.475)	0.154	0.12	(0.651)
Midwife in the village	0.411	0.409	(0.026)	0.409	0.356	(0.682)
Majority has access to piped water	0.088	0.145	(-1.002)	0.089	0.123	(-0.601)
Main road is paved	0.604	0.503	(1.418)	0.63	0.459	(2.291) **
Majority traffic through land	0.883	0.972	(-2.064) **	0.872	0.968	(-1.958) *
Bus terminal in the village	0.033	0.035	(-0.111)	0.036	0.023	(1.208)
Post office in village	0.073	0.075	(-0.084)	0.074	0.067	(0.275)
Land area (ha)	43567.841	16166.957	(2.588) ***	44470.253	19744.98	(1.967) **
Ratio population/land (ha)	0.862	2.076	(-1.143)	0.974	0.892	(0.131)
Permanent market in village	0.121	0.109	(0.330)	0.121	0.098	(0.582)
Temporary market in village	0.14	0.104	(1.217)	0.134	0.092	(1.362)
Community safety post	0.877	0.703	(1.373)	0.865	0.66	(1.381)
Police house in village	0.1	0.067	(0.922)	0.107	0.059	(1.196)
Village head finished junior-high	0.714	0.806	(-1.390)	0.713	0.772	(-0.837)
Village head finished high school	0.427	0.458	(-0.409)	0.445	0.417	(0.356)
Number of villages	*2308*	*2799*		*1949*	*2160*	
Number of districts	*12*	*14*		*12*	*11*	

Note: the "Difference" columns report the difference-in-mean test between oil-rich and oil-scarce villages, where standard errors have been clustered at the district level using district borders as in 1990

TABLE 2B

DESCRIPTIVE STATISTICS EXCLUDING PRODUCING DISTRICTS: SUMATRA AREA

Sample	<300 kilometers			<200 kilometers		
Treatment	Oil-rich	Oil-scarce	Difference	Oil-rich	Oil-scarce	Difference
Statistic	mean	mean	t-stat	mean	mean	t-stat
Village located on the coast	0.033	0.031	(0.080)	0.039	0.029	(0.342)
Village located in a valley	0.027	0.081	(-1.339)	0.026	0.097	(-1.701) *
Village located in a hilly area	0.122	0.132	(-0.160)	0.088	0.135	(-0.815)
Village located in a plane	0.822	0.756	(0.796)	0.852	0.739	(1.317)
Urban village	0.106	0.19	(-0.915)	0.12	0.122	(-0.024)
Official urban village	0.123	0.166	(-0.459)	0.145	0.118	(0.352)
Population	3106.973	3045.089	(0.053)	2679.367	2419.133	(0.259)
Number of households	542.235	572.193	(-0.147)	553.303	440.103	(0.596)
Number of households in agriculture	285.767	244.298	(0.579)	266.586	249.06	(0.206)
Electricity in the village	0.856	0.89	(-0.651)	0.878	0.875	(0.054)
Share of households with electricity	0.337	0.538	(-2.291) **	0.366	0.501	(-1.476)
Public phone in the village	0.103	0.143	(-0.496)	0.113	0.084	(0.443)
Number of households with phone[1]	0.021	0.023	(-0.135)	0.024	0.015	(0.746)
Main road is lighted	0.274	0.45	(-1.445)	0.309	0.377	(-0.605)
Majority uses LPG/Kerosene for cooking	0.125	0.199	(-0.819)	0.138	0.12	(0.250)
Majority litters in bin, then delivered	0.132	0.139	(-0.100)	0.149	0.105	(0.684)
Majority households has private toilet	0.356	0.457	(-0.615)	0.342	0.392	(-0.293)
Majority households has public toilet	0.592	0.44	(1.265)	0.608	0.488	(0.943)
Streaming sewage system in village	0.59	0.67	(-0.911)	0.574	0.651	(-0.721)
Share of permanent dwellings	0.147	0.194	(-0.825)	0.143	0.158	(-0.290)
Slum in village	0.128	0.059	(1.821) *	0.139	0.049	(2.235) **
Share of households living in slum	0.018	0.007	(2.269) **	0.02	0.005	(2.696) ***
Primary school in the village	0.9	0.799	(1.029)	0.893	0.789	(0.896)
Junior-high school in the village	0.263	0.328	(-0.738)	0.266	0.286	(-0.219)
Senior-high school in the village	0.122	0.178	(-0.944)	0.122	0.136	(-0.273)
Number of primary schools in the village[1]	2.31	2.766	(-0.562)	2.313	2.609	(-0.320)
Number of junior-high schools in the village[1]	0.418	0.59	(-0.856)	0.428	0.462	(-0.188)
Number of senior-high schools in the village[1]	0.208	0.353	(-1.071)	0.206	0.25	(-0.430)
Hospital in the village	0.018	0.046	(-1.219)	0.02	0.028	(-0.507)
Maternity house in the village	0.051	0.132	(-1.292)	0.057	0.071	(-0.317)
Health center in the village	0.094	0.102	(-0.241)	0.098	0.091	(0.182)
Doctor in the village	0.149	0.179	(-0.387)	0.159	0.118	(0.680)
Midwife in the village	0.445	0.407	(0.437)	0.446	0.358	(1.032)
Majority has access to piped water	0.101	0.15	(-0.736)	0.104	0.125	(-0.336)
Main road is paved	0.625	0.509	(1.429)	0.66	0.464	(2.378) **
Majority traffic through land	0.911	0.981	(-1.657) *	0.897	0.977	(-1.605)
Bus terminal in the village	0.03	0.033	(-0.235)	0.033	0.024	(0.778)
Post office in village	0.075	0.073	(0.065)	0.077	0.066	(0.407)
Land area (ha)	34759.007	15471.433	(2.196) **	34550.818	17741.619	(1.565)
Ratio population/land (ha)	0.998	1.708	(-0.674)	1.157	0.916	(0.307)
Permanent market in village	0.121	0.109	(0.298)	0.119	0.094	(0.587)
Temporary market in village	0.14	0.099	(1.280)	0.13	0.092	(1.121)
Community safety post	0.893	0.676	(1.629)	0.88	0.65	(1.520)
Police house in village	0.088	0.064	(0.639)	0.096	0.053	(0.990)
Village head finished junior-high	0.717	0.801	(-1.166)	0.709	0.772	(-0.841)
Village head finished high school	0.428	0.46	(-0.378)	0.448	0.417	(0.362)
Number of villages	*1825*	*2547*		*1497*	*2097*	
Number of districts	*11*	*13*		*11*	*11*	

Note: the "Difference" columns report the difference-in-mean test between oil-rich and oil-scarce villages, where standard errors have been clustered at the district level using district borders as in 1990

TABLE 3A

IMPACT OF RESOURCE TRANSFERS ON TRANSPORT INFRASTRUCTURES: SUMATRA AREA

	Majority traffic through land				Access to paved road				Access to bus terminal			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PANEL A: BINARY TREATMENT												
TREATMENT × 2005	0.057*	0.057	0.001	0.011	0.027	0.031	0.024	0.031	-0.001	-0.002	-0.016*	-0.016*
	(0.030)	(0.034)	(0.041)	(0.040)	(0.026)	(0.025)	(0.046)	(0.045)	(0.009)	(0.007)	(0.008)	(0.009)
TREATMENT × 2002	0.025**	0.024*	0.012	0.016	0.027	0.019	0.057	0.064	-0.010	0.001	-0.004	-0.002
	(0.011)	(0.013)	(0.021)	(0.021)	(0.028)	(0.031)	(0.057)	(0.052)	(0.011)	(0.007)	(0.018)	(0.019)
TREATMENT × 1999	0.020*	0.021*	0.021	0.027*	-0.033	-0.067	-0.057*	-0.048*	-0.005	-0.003	-0.017	-0.017
	(0.010)	(0.012)	(0.014)	(0.015)	(0.048)	(0.051)	(0.033)	(0.025)	(0.006)	(0.006)	(0.011)	(0.012)
R-squared	0.030	0.033	0.054	0.057	0.004	0.006	0.008	0.008	0.001	0.000	0.001	0.002
PANEL B: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS	0.001***	0.001**	0.000	0.000	0.001***	0.002***	0.002***	0.002***	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.033	0.032	0.054	0.057	0.004	0.005	0.009	0.010	0.001	0.000	0.001	0.001
PANEL C: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS × 2005	0.002***	0.001**	0.000	0.001	0.001**	0.001***	0.001***	0.001***	-0.000	-0.000**	-0.000*	-0.000*
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
PER CAPITA OIL-GAS × 2002	0.001**	0.001***	0.000***	0.000***	0.003***	0.003***	0.004***	0.004***	0.000	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.035	0.033	0.054	0.057	0.006	0.007	0.011	0.012	0.001	0.002	0.002	0.003
Village FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE or Segment-year FE	year	year	segment	segment	year	year	segment	segment	year	year	segment	segment
Sample	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km
Geographical controls				cubic				cubic				cubic
Observations	20,423	16,431	16,431	16,431	19,410	15,511	15,511	15,511	20,423	16,431	16,431	16,431
Number of vv	5,107	4,109	4,109	4,109	5,025	4,033	4,033	4,033	5,107	4,109	4,109	4,109
Number of clusters	26	23	23	23	26	23	23	23	26	23	23	23

Note: robust standard errors in parentheses clustered at the district level as in 1993. *** p<0.01, ** p<0.05, * p<0.1. All specifications exclude mountainous areas, i.e., parts of North Sumatra and Jambi. Columns (1)-(3), (5)-(7), (9)-(11) show the coefficient estimates associated with specification (1), i.e., the Difference-in-Difference strategy. Columns (4), (8) and (12) show the coefficient estimates associated with specification (2), i.e., the spatial Regression Discontinuity strategy.

TABLE 3B

IMPACT OF RESOURCE TRANSFERS ON EDUCATION INFRASTRUCTURES: SUMATRA AREA

	Primary school in village				Junior high school				Senior high school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PANEL A: BINARY TREATMENT												
TREATMENT × 2005	0.010	0.014*	0.012	0.013	0.062**	0.069***	0.090**	0.099**	0.032*	0.037*	0.039	0.045
	(0.007)	(0.008)	(0.010)	(0.009)	(0.023)	(0.024)	(0.043)	(0.039)	(0.017)	(0.020)	(0.041)	(0.041)
TREATMENT × 2002	0.008	0.013**	0.002	0.004	0.049**	0.058***	0.079**	0.083**	0.015	0.018	0.007	0.014
	(0.006)	(0.006)	(0.007)	(0.007)	(0.019)	(0.020)	(0.031)	(0.030)	(0.013)	(0.015)	(0.019)	(0.019)
TREATMENT × 1999	0.006	0.011*	0.008	0.010	0.022	0.023	0.018	0.022	0.010	0.006	-0.007	-0.002
	(0.006)	(0.006)	(0.011)	(0.011)	(0.016)	(0.018)	(0.024)	(0.022)	(0.008)	(0.010)	(0.015)	(0.014)
R-squared	0.001	0.001	0.002	0.002	0.016	0.017	0.019	0.020	0.015	0.017	0.019	0.020
PANEL B: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS	-0.000	-0.000	0.000	0.000	0.002***	0.002***	0.002***	0.002***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.001	0.001	0.002	0.002	0.020	0.020	0.021	0.021	0.016	0.019	0.020	0.021
PANEL C: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS × 2005	0.000	0.000	0.000	0.000	0.002***	0.002***	0.002***	0.002***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
PER CAPITA OIL-GAS × 2002	-0.000**	-0.000*	-0.000***	-0.000***	0.002***	0.002***	0.002***	0.002***	0.001***	0.001***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.001	0.001	0.002	0.002	0.020	0.020	0.021	0.021	0.016	0.019	0.021	0.022
Village FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE or Segment-year FE	year	year	segment	segment	year	year	segment	segment	year	year	segment	segment
Sample	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km
Geographical controls				cubic				cubic				cubic
Observations	20,423	16,431	16,431	16,431	20,423	16,431	16,431	16,431	20,423	16,431	16,431	16,431
Number of vv	5,107	4,109	4,109	4,109	5,107	4,109	4,109	4,109	5,107	4,109	4,109	4,109
Number of clusters	26	23	23	23	26	23	23	23	26	23	23	23

Note: robust standard errors in parentheses clustered at the district level as in 1993. *** p<0.01, ** p<0.05, * p<0.1. All specifications exclude mountainous areas, i.e., parts of North Sumatra and Jambi. Columns (1)-(3), (5)-(7), (9)-(11) show the coefficient estimates associated with specification (1), i.e., the Difference-in-Difference strategy. Columns (4), (8) and (12) show the coefficient estimates associated with specification (2), i.e., the spatial Regression Discontinuity strategy.

TABLE 3C

IMPACT OF RESOURCE TRANSFERS ON HEALTH INFRASTRUCTURES: SUMATRA AREA

	Maternity hospital				Health center				Access to clean water			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PANEL A: BINARY TREATMENT												
TREATMENT × 2005	0.031	0.018	0.026	0.029	0.014*	0.013	0.016	0.017	-0.017	-0.008	-0.003	0.002
	(0.021)	(0.020)	(0.027)	(0.026)	(0.007)	(0.009)	(0.016)	(0.016)	(0.014)	(0.013)	(0.016)	(0.015)
TREATMENT × 2002	0.023	0.012	0.008	0.010	0.007	0.004	0.013	0.009	-0.031**	-0.020*	-0.031	-0.027
	(0.016)	(0.016)	(0.023)	(0.022)	(0.008)	(0.009)	(0.014)	(0.015)	(0.013)	(0.011)	(0.020)	(0.018)
TREATMENT × 1999	0.020*	0.016	0.024*	0.023	-0.003	0.001	0.005	0.005	0.010	0.011	0.000	0.003
	(0.011)	(0.012)	(0.013)	(0.014)	(0.006)	(0.007)	(0.009)	(0.010)	(0.011)	(0.010)	(0.013)	(0.012)
R-squared	0.002	0.001	0.002	0.002	0.002	0.003	0.003	0.004	0.003	0.002	0.004	0.004
PANEL B: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS	0.001***	0.001***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.002	0.002	0.002	0.003	0.002	0.003	0.004	0.004	0.000	0.001	0.003	0.003
PANEL C: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS × 2005	0.001**	0.001**	0.000**	0.000***	0.000***	0.000***	0.000***	0.000***	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
PER CAPITA OIL-GAS × 2002	0.001***	0.000***	0.000**	0.000**	0.000***	0.000**	0.000***	0.000***	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.002	0.002	0.002	0.003	0.002	0.003	0.004	0.004	0.000	0.001	0.003	0.003
Village FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE or Segment-year FE	year	year	segment	segment	year	year	segment	segment	year	year	segment	segment
Sample	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km
Geographical controls				cubic				cubic				cubic
Observations	20,423	16,431	16,431	16,431	20,423	16,431	16,431	16,431	20,423	16,431	16,431	16,431
Number of vv	5,107	4,109	4,109	4,109	5,107	4,109	4,109	4,109	5,107	4,109	4,109	4,109
Number of clusters	26	23	23	23	26	23	23	23	26	23	23	23

Note: robust standard errors in parentheses clustered at the district level as in 1993. *** p<0.01, ** p<0.05, * p<0.1. All specifications exclude mountainous areas, i.e., parts of North Sumatra and Jambi. Columns (1)-(3), (5)-(7), (9)-(11) show the coefficient estimates associated with specification (1), i.e., the Difference-in-Difference strategy. Columns (4), (8) and (12) show the coefficient estimates associated with specification (2), i.e., the spatial Regression Discontinuity strategy.

TABLE 3D

IMPACT OF RESOURCE TRANSFERS ON OTHER INFRASTRUCTURES: SUMATRA AREA

	Public phone in village				Post office in village				Permanent market in village			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PANEL A: BINARY TREATMENT												
TREATMENT × 2005	-0.061	-0.044	-0.012	-0.031	0.004	0.006	0.018	0.020	0.022	0.037*	0.075***	0.079***
	(0.058)	(0.063)	(0.046)	(0.041)	(0.008)	(0.009)	(0.012)	(0.013)	(0.022)	(0.021)	(0.019)	(0.020)
TREATMENT × 2002	-0.017	0.003	0.040	0.033	0.008	0.009	0.020	0.023*	0.036	0.053**	0.070**	0.077**
	(0.041)	(0.046)	(0.040)	(0.039)	(0.006)	(0.006)	(0.012)	(0.012)	(0.022)	(0.022)	(0.031)	(0.032)
TREATMENT × 1999	-0.013	-0.017	0.001	0.001	0.004	0.005	0.010	0.012	0.003	0.012	0.019	0.024
	(0.011)	(0.013)	(0.012)	(0.011)	(0.006)	(0.007)	(0.007)	(0.008)	(0.017)	(0.015)	(0.015)	(0.015)
R-squared	0.156	0.146	0.148	0.152	0.001	0.001	0.002	0.003	0.008	0.011	0.014	0.015
PANEL B: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS	0.000	0.000	0.000	0.000	0.000***	0.000**	0.000***	0.000***	0.001***	0.001***	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.154	0.145	0.147	0.151	0.001	0.002	0.002	0.003	0.010	0.013	0.014	0.015
PANEL C: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS × 2005	-0.000	-0.000	-0.000	-0.000	0.000**	0.000**	0.000***	0.000***	0.001***	0.001***	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
PER CAPITA OIL-GAS × 2002	0.001**	0.002**	0.002**	0.002**	0.000***	0.001***	0.001***	0.001***	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.155	0.147	0.149	0.152	0.002	0.002	0.003	0.003	0.011	0.014	0.015	0.016
Village FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE or Segment-year FE	year	year	segment	segment	year	year	segment	segment	year	year	segment	segment
Sample	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km
Geographical controls				cubic				cubic				cubic
Observations	20,423	16,431	16,431	16,431	20,423	16,431	16,431	16,431	20,422	16,430	16,430	16,430
Number of vv	5,107	4,109	4,109	4,109	5,107	4,109	4,109	4,109	5,107	4,109	4,109	4,109
Number of clusters	26	23	23	23	26	23	23	23	26	23	23	23

Note: robust standard errors in parentheses clustered at the district level as in 1993. *** p<0.01, ** p<0.05, * p<0.1. All specifications exclude mountainous areas, i.e., parts of North Sumatra and Jambi. Columns (1)-(3), (5)-(7), (9)-(11) show the coefficient estimates associated with specification (1), i.e., the Difference-in-Difference strategy. Columns (4), (8) and (12) show the coefficient estimates associated with specification (2), i.e., the spatial Regression Discontinuity strategy.

TABLE 4

IMPACT OF RESOURCE TRANSFERS CONTROLLING FOR PRODUCING DISTRICTS: SUMATRA AREA

	Access to paved road				Junior high school				Permanent market in village			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PANEL A: BINARY TREATMENT												
TREATMENT × 2005	0.031	0.034	0.028	0.037	0.054***	0.056**	0.071*	0.081**	0.014	0.025	0.059***	0.062***
	(0.025)	(0.024)	(0.047)	(0.045)	(0.018)	(0.022)	(0.036)	(0.033)	(0.018)	(0.018)	(0.011)	(0.012)
TREATMENT × 2002	0.026	0.005	0.038	0.046	0.047**	0.058***	0.081**	0.087***	0.027	0.034**	0.042**	0.048**
	(0.025)	(0.021)	(0.045)	(0.040)	(0.017)	(0.019)	(0.032)	(0.030)	(0.017)	(0.015)	(0.015)	(0.018)
TREATMENT × 1999	-0.025	-0.056	-0.036	-0.021	0.020	0.024	0.018	0.025	-0.004	-0.002	-0.003	0.000
	(0.045)	(0.050)	(0.028)	(0.016)	(0.016)	(0.019)	(0.026)	(0.024)	(0.016)	(0.011)	(0.013)	(0.013)
PRODUCING DISTRICT × 2005	-0.037	-0.013	-0.011	-0.018	0.064*	0.063*	0.060**	0.052	0.071***	0.060**	0.052***	0.050***
	(0.037)	(0.044)	(0.045)	(0.046)	(0.032)	(0.037)	(0.029)	(0.031)	(0.019)	(0.026)	(0.010)	(0.011)
PRODUCING DISTRICT × 2002	0.011	0.071	0.058	0.053	0.016	-0.000	-0.007	-0.012	0.076***	0.091***	0.091***	0.086***
	(0.067)	(0.080)	(0.067)	(0.069)	(0.033)	(0.042)	(0.034)	(0.036)	(0.020)	(0.019)	(0.017)	(0.016)
PRODUCING DISTRICT × 1999	-0.076**	-0.056**	-0.063***	-0.075**	0.010	-0.005	-0.003	-0.008	0.053**	0.068***	0.070***	0.067***
	(0.035)	(0.023)	(0.018)	(0.027)	(0.016)	(0.016)	(0.016)	(0.018)	(0.022)	(0.020)	(0.023)	(0.022)
R-squared	0.006	0.008	0.010	0.011	0.018	0.019	0.020	0.021	0.011	0.014	0.017	0.018
PANEL B: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS	0.002***	0.002***	0.003***	0.003***	0.002***	0.002***	0.002***	0.002***	0.001***	0.002***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
PRODUCING DISTRICT × 2005	-0.084***	-0.087***	-0.114***	-0.117***	0.014	0.007	-0.000	-0.006	0.035	0.006	0.016	0.015
	(0.019)	(0.029)	(0.026)	(0.027)	(0.017)	(0.020)	(0.023)	(0.024)	(0.023)	(0.022)	(0.031)	(0.034)
PRODUCING DISTRICT × 2002	-0.021	0.008	-0.008	-0.010	-0.018	-0.032	-0.038	-0.042	0.054**	0.060**	0.065**	0.063**
	(0.050)	(0.060)	(0.048)	(0.050)	(0.020)	(0.022)	(0.026)	(0.026)	(0.021)	(0.029)	(0.029)	(0.030)
PRODUCING DISTRICT × 1999	-0.082**	-0.085**	-0.079***	-0.085***	0.014	0.006	0.004	0.002	0.052**	0.067***	0.069***	0.067***
	(0.038)	(0.039)	(0.027)	(0.028)	(0.015)	(0.016)	(0.019)	(0.021)	(0.020)	(0.020)	(0.023)	(0.022)
R-squared	0.006	0.008	0.012	0.013	0.021	0.021	0.021	0.022	0.012	0.016	0.017	0.017
PANEL C: CONTINUOUS TREATMENT												
PER CAPITA OIL-GAS × 2005	0.001***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
PER CAPITA OIL-GAS × 2002	0.003***	0.003***	0.004***	0.004***	0.002***	0.002***	0.003***	0.003***	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
PRODUCING DISTRICT × 2005	-0.069***	-0.069**	-0.090***	-0.094***	0.016	0.012	0.011	0.005	0.041	0.011	0.023	0.022
	(0.023)	(0.029)	(0.031)	(0.032)	(0.017)	(0.021)	(0.026)	(0.026)	(0.024)	(0.023)	(0.031)	(0.033)
PRODUCING DISTRICT × 2002	-0.050	-0.026	-0.050	-0.051	-0.021	-0.041	-0.057**	-0.060**	0.044*	0.051	0.055	0.050
	(0.038)	(0.057)	(0.041)	(0.044)	(0.023)	(0.025)	(0.027)	(0.029)	(0.025)	(0.036)	(0.035)	(0.036)
PRODUCING DISTRICT × 1999	-0.082**	-0.084**	-0.079***	-0.085***	0.014	0.006	0.004	0.002	0.052**	0.067***	0.069***	0.067***
	(0.038)	(0.039)	(0.027)	(0.028)	(0.015)	(0.016)	(0.019)	(0.021)	(0.020)	(0.020)	(0.023)	(0.022)
R-squared	0.007	0.008	0.012	0.013	0.021	0.021	0.022	0.022	0.012	0.016	0.017	0.017
Village FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE or Segment-year FE	year	year	segment	segment	year	year	segment	segment	year	year	segment	segment
Sample	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km	<300km	<200km	<200km	<200km
Geographical controls				cubic				cubic				cubic
Observations	19,410	15,511	15,511	15,511	20,423	16,431	16,431	16,431	20,422	16,430	16,430	16,430
Number of vv	5,025	4,033	4,033	4,033	5,107	4,109	4,109	4,109	5,107	4,109	4,109	4,109
Number of clusters	26	23	23	23	26	23	23	23	26	23	23	23

Note: robust standard errors in parentheses clustered at the district level as in 1993. *** p<0.01, ** p<0.05, * p<0.1. All specifications exclude mountainous areas, i.e., parts of North Sumatra and Jambi. Columns (1)-(3), (5)-(7), (9)-(11) show the coefficient estimates associated with specification (1), i.e., the Difference-in-Difference strategy. Columns (4), (8) and (12) show the coefficient estimates associated with specification (2), i.e., the spatial Regression Discontinuity strategy.

TABLE 5

DESCRIPTIVE STATISTICS: KALIMANTAN AREA

Sample	<100 kilometers			<75 kilometers			<50 kilometers			<25 kilometers		
Treatment	Oil-rich	Oil-scarce	Difference	Oil-rich	Oil-scarce	Difference	Oil-rich	Oil-scarce	Difference	Oil-rich	Oil-scarce	Difference
Statistic	mean	mean	t-stat	mean	mean	t-stat	mean	mean	t-stat	mean	mean	t-stat
Village located on the coast	0.065	0.033	(0.518)	0.045	0.028	(0.376)	0.019	0.039	(-0.680)	0.019	0.059	(-0.953)
Village located in a valley	0.058	0.062	(-0.098)	0.068	0.066	(0.035)	0.076	0.111	(-0.602)	0.019	0.141	(-1.401)
Village located in a hilly area	0.178	0.115	(0.752)	0.194	0.122	(0.707)	0.209	0.114	(0.892)	0.327	0.119	(1.316)
Village located in a plane	0.698	0.8	(-0.794)	0.694	0.796	(-0.801)	0.696	0.749	(-0.459)	0.635	0.681	(-0.275)
Urban village	0.011	0.031	(-4.391) ***	0.009	0.024	(-2.965) ***	0.013	0.005	(1.385)	0	0	.
Official urban village	0.069	0.035	(0.527)	0.023	0.028	(-0.254)	0.013	0.016	(-0.198)	0	0	.
Population	986.135	1066.572	(-0.226)	877.896	978.426	(-0.301)	817.563	847.271	(-0.095)	773.385	805.844	(-0.155)
Number of households	220.095	249.574	(-0.347)	201.396	232.907	(-0.383)	184.114	206.914	(-0.287)	173.346	198.696	(-0.412)
Number of households in agriculture	142.789	138.293	(0.100)	132.257	131.487	(0.018)	111.253	139.278	(-0.870)	110.154	150.993	(-0.967)
Electricity in the village	0.625	0.81	(-2.654) ***	0.595	0.817	(-2.960) ***	0.627	0.752	(-1.175)	0.596	0.689	(-0.381)
Share of households with electricity	0.287	0.394	(-1.964) **	0.267	0.395	(-2.182) **	0.294	0.317	(-0.301)	0.269	0.22	(0.359)
Public phone in the village	0.011	0.043	(-2.774) ***	0.014	0.037	(-2.105) **	0.013	0.026	(-0.549)	0	0.007	(-1.089)
Number of households with phone[1]	0.002	0.007	(-3.148) ***	0.002	0.006	(-1.781) *	0.002	0.002	(0.070)	0	0	(-1.089)
Main road is lighted	0.196	0.498	(-2.891) ***	0.171	0.485	(-2.699) ***	0.171	0.394	(-2.102) **	0.154	0.156	(-0.023)
Majority uses LPG/Kerosene for cooking	0.033	0.024	(0.398)	0.023	0.021	(0.142)	0.025	0.002	(1.353)	0.019	0	(1.116)
Majority litters in bin, then delivered	0.015	0.048	(-2.769) ***	0.009	0.046	(-2.628) ***	0.013	0.021	(-0.467)	0	0	.
Majority households has private toilet	0.135	0.165	(-0.379)	0.131	0.159	(-0.358)	0.095	0.114	(-0.262)	0.038	0.133	(-1.355)
Majority households has public toilet	0.793	0.803	(-0.161)	0.797	0.799	(-0.039)	0.816	0.821	(-0.067)	0.885	0.793	(0.978)
Streaming sewage system in village	0.204	0.342	(-2.256) **	0.203	0.351	(-1.789) *	0.184	0.411	(-2.012) **	0.058	0.437	(-1.712) *
Share of permanent dwellings	0.028	0.02	(0.778)	0.019	0.02	(-0.144)	0.017	0.017	(-0.011)	0.004	0.015	(-0.884)
Slum in village	0.036	0.042	(-0.185)	0.041	0.015	(0.910)	0.038	0.007	(1.472)	0.077	0.007	(1.745) *
Share of households living in slum	0.011	0.01	(0.199)	0.012	0.003	(1.400)	0.014	0.002	(1.987) **	0.009	0.006	(0.382)

Note: the "Difference" columns report the difference-in-mean test between oil-rich and oil-scarce villages, where standard errors have been clustered at the district level using district borders as in 1990

TABLE 5 (CONTINUED)

DESCRIPTIVE STATISTICS: KALIMANTAN AREA

Sample	<100 kilometers			<75 kilometers			<50 kilometers			<25 kilometers		
Treatment	Oil-rich	Oil-scarce	Difference	Oil-rich	Oil-scarce	Difference	Oil-rich	Oil-scarce	Difference	Oil-rich	Oil-scarce	Difference
Statistic	mean	mean	t-stat	mean	mean	t-stat	mean	mean	t-stat	mean	mean	t-stat
Primary school in the village	0.825	0.927	(-1.435)	0.802	0.92	(-1.488)	0.785	0.926	(-1.636)	0.788	0.881	(-1.050)
Junior-high school in the village	0.156	0.165	(-0.172)	0.135	0.152	(-0.332)	0.139	0.139	(0.001)	0.135	0.141	(-0.135)
Senior-high school in the village	0.044	0.051	(-0.781)	0.041	0.047	(-0.631)	0.044	0.035	(0.579)	0	0.022	(-1.089)
Number of primary schools in the village[1]	1.407	1.779	(-0.779)	1.351	1.742	(-0.759)	1.247	1.626	(-0.884)	1.019	1.496	(-1.778) *
Number of junior-high schools in the village[1]	0.193	0.203	(-0.127)	0.176	0.185	(-0.108)	0.177	0.158	(0.239)	0.135	0.156	(-0.367)
Number of senior-high schools in the village[1]	0.062	0.071	(-0.390)	0.063	0.064	(-0.033)	0.07	0.039	(0.773)	0	0.03	(-1.089)
Hospital in the village	0.007	0.008	(-0.212)	0.009	0.007	(0.509)	0.013	0.005	(1.385)	0	0	.
Maternity house in the village	0.004	0.005	(-0.256)	0.005	0.003	(0.272)	0.006	0.005	(0.236)	0	0	.
Health center in the village	0.102	0.079	(1.190)	0.086	0.077	(0.315)	0.089	0.07	(1.010)	0.115	0.089	(1.398)
Doctor in the village	0.113	0.079	(1.320)	0.108	0.076	(0.997)	0.12	0.072	(1.621)	0.115	0.074	(1.245)
Midwife in the village	0.251	0.168	(0.919)	0.243	0.153	(0.838)	0.209	0.128	(0.858)	0.192	0.119	(0.981)
Majority has access to piped water	0.062	0.166	(-3.220) ***	0.032	0.152	(-2.784) ***	0.025	0.093	(-1.901) *	0.019	0.015	(0.201)
Main road is paved	0.283	0.659	(-2.518) **	0.287	0.631	(-2.294) **	0.321	0.628	(-1.906) *	0.35	0.441	(-0.471)
Majority traffic through road	0.502	0.823	(-3.118) ***	0.518	0.866	(-3.539) ***	0.494	0.805	(-2.140) **	0.385	0.756	(-2.340) **
Bus terminal in the village	0.022	0.025	(-0.303)	0.023	0.025	(-0.272)	0.019	0.012	(0.539)	0	0.007	(-1.089)
Post office in village	0.084	0.063	(1.924) *	0.072	0.058	(1.107)	0.076	0.053	(1.565)	0.077	0.067	(0.384)
Land area (ha)	46528.305	31434.728	(0.874)	48882.959	28070.428	(1.141)	54919.823	37296.427	(0.778)	66531.25	59518	(0.165)
Ratio population/land (ha)	0.084	0.235	(-2.157) **	0.084	0.22	(-1.656) *	0.098	0.08	(0.430)	0.16	0.035	(4.999) ***
Permanent market in village	0.08	0.091	(-0.270)	0.072	0.104	(-0.900)	0.063	0.132	(-1.896) *	0.077	0.133	(-0.825)
Temporary market in village	0.091	0.096	(-0.173)	0.108	0.099	(0.272)	0.127	0.1	(0.717)	0.231	0.141	(1.232)
Community safety post	0.64	0.752	(-1.270)	0.653	0.73	(-0.682)	0.62	0.766	(-0.942)	0.442	0.652	(-1.028)
Police house in village	0.047	0.06	(-0.523)	0.041	0.056	(-0.641)	0.032	0.06	(-1.173)	0.019	0.052	(-1.186)
Village head finished junior-high	0.349	0.473	(-1.750) *	0.333	0.457	(-1.551)	0.354	0.397	(-0.424)	0.288	0.348	(-0.435)
Village head finished high school	0.178	0.218	(-0.694)	0.144	0.211	(-1.491)	0.177	0.165	(0.240)	0.135	0.156	(-0.279)
Number of villages	*275*	*1276*		*222*	*952*		*158*	*431*		*52*	*135*	
Number of districts	*4*	*11*		*4*	*10*		*2*	*6*		*2*	*4*	

Note: the "Difference" columns report the difference-in-mean test between oil-rich and oil-scarce villages, where standard errors have been clustered at the district level using district borders as in 1990

TABLE 6A

IMPACT OF RESOURCE TRANSFERS ON TRANSPORTATION AND EDUCATION INFRASTRUCTURES: KALIMANTAN AREA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
PANEL A: TRANSPORTATION INFRASTRUCTURES															
	Majority traffic through land					Main road is paved					Bus terminal				
TREATMENT × 2005	0.242***	0.267***	0.246**	0.236**	0.235**	0.070**	0.044	-0.039	0.005	0.008	0.002	0.010	0.010	0.010*	0.010*
	(0.037)	(0.047)	(0.084)	(0.091)	(0.085)	(0.025)	(0.025)	(0.052)	(0.032)	(0.031)	(0.010)	(0.008)	(0.006)	(0.004)	(0.004)
TREATMENT × 2002	0.117***	0.146***	0.116	0.151**	0.150**	0.007	-0.045**	-0.113	-0.068	-0.068	-0.005	0.001	0.001	0.003	0.003
	(0.036)	(0.035)	(0.072)	(0.057)	(0.055)	(0.035)	(0.018)	(0.061)	(0.040)	(0.040)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
TREATMENT × 1999	-0.062	-0.062	-0.071	-0.009	-0.009	-0.035**	-0.060***	-0.095*	-0.051*	-0.048*	0.004	0.011**	0.023*	0.029***	0.029***
	(0.074)	(0.076)	(0.085)	(0.074)	(0.074)	(0.014)	(0.016)	(0.047)	(0.024)	(0.022)	(0.005)	(0.005)	(0.010)	(0.008)	(0.008)
Observations	6,204	4,696	2,356	2,356	2,356	4,997	3,942	1,813	1,813	1,813	6,204	4,696	2,356	2,356	2,356
R-squared	0.112	0.131	0.139	0.154	0.163	0.007	0.014	0.022	0.024	0.029	0.001	0.001	0.005	0.006	0.006
Number of villages	1,551	1,174	589	589	589	1,376	1,081	523	523	523	1,551	1,174	589	589	589
Number of clusters	15	14	8	8	8	15	13	8	8	8	15	14	8	8	8
PANEL A: EDUCATION INFRASTRUCTURES															
	Primary school in village					Junior high school					Senior high school				
TREATMENT × 2005	0.082**	0.090*	0.115*	0.088	0.087	0.119***	0.113***	0.111**	0.094**	0.093**	0.127***	0.102***	0.113***	0.107**	0.106**
	(0.037)	(0.044)	(0.050)	(0.056)	(0.057)	(0.017)	(0.035)	(0.033)	(0.028)	(0.030)	(0.012)	(0.022)	(0.032)	(0.036)	(0.040)
TREATMENT × 2002	0.015	-0.001	-0.005	-0.023	-0.023	0.056***	0.042	0.039*	0.028*	0.028*	0.025	0.013	0.008	0.023*	0.022**
	(0.015)	(0.014)	(0.013)	(0.014)	(0.015)	(0.008)	(0.024)	(0.019)	(0.015)	(0.014)	(0.023)	(0.013)	(0.015)	(0.011)	(0.009)
TREATMENT × 1999	0.033*	0.034	0.029	0.024	0.024	0.036	0.041	0.038	0.008	0.007	0.019	0.007*	0.003	0.005	0.005
	(0.017)	(0.020)	(0.018)	(0.021)	(0.021)	(0.034)	(0.049)	(0.049)	(0.045)	(0.047)	(0.018)	(0.004)	(0.007)	(0.010)	(0.010)
Observations	6,204	4,696	2,356	2,356	2,356	6,204	4,696	2,356	2,356	2,356	6,204	4,696	2,356	2,356	2,356
R-squared	0.016	0.024	0.035	0.046	0.048	0.036	0.035	0.051	0.057	0.060	0.044	0.040	0.063	0.076	0.086
Number of villages	1,551	1,174	589	589	589	1,551	1,174	589	589	589	1,551	1,174	589	589	589
Number of clusters	15	14	8	8	8	15	14	8	8	8	15	14	8	8	8
Village FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE or Segment-year FE	year	year	year	segment	segment	year	year	year	segment	segment	year	year	year	segment	segment
Sample	<100km	<75km	<50km	<50km	<50km	<100km	<75km	<50km	<50km	<50km	<100km	<75km	<50km	<50km	<50km
Geographical controls					cubic					cubic					cubic

Note: robust standard errors in parentheses clustered at the district level as in 1993. *** p<0.01, ** p<0.05, * p<0.1. Columns (1)-(3), (6)-(8), (11)-(13) show the coefficient estimates associated with specification (1), i.e., the Difference-in-Difference strategy. Columns (4)-(5), (9)-(10) and (14)-(15) show the coefficient estimates associated with specification (2), i.e., the spatial Regression Discontinuity strategy.

TABLE 6B

IMPACT OF RESOURCE TRANSFERS ON HEALTH AND OTHER INFRASTRUCTURES: KALIMANTAN AREA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
PANEL A: HEALTH INFRASTRUCTURES															
	Maternity house in village					Health center in village					Access to piped water				
TREATMENT × 2005	0.013***	0.015***	0.019***	0.026***	0.026***	0.073*	0.090*	0.110**	0.086*	0.086*	-0.039	-0.038	-0.057	-0.024	-0.023
	(0.003)	(0.002)	(0.004)	(0.004)	(0.005)	(0.038)	(0.048)	(0.044)	(0.041)	(0.043)	(0.028)	(0.022)	(0.036)	(0.040)	(0.037)
TREATMENT × 2002	0.008	0.007	0.015	0.028**	0.028**	-0.019***	-0.013**	-0.014	-0.008	-0.008	-0.019	0.001	0.007	0.031*	0.032**
	(0.007)	(0.009)	(0.013)	(0.010)	(0.010)	(0.006)	(0.005)	(0.009)	(0.007)	(0.007)	(0.021)	(0.016)	(0.021)	(0.014)	(0.013)
TREATMENT × 1999	0.006**	0.005	0.011**	0.013*	0.013*	-0.021***	-0.013***	-0.014*	-0.008	-0.008	-0.015	-0.008	-0.021*	-0.012	-0.011
	(0.002)	(0.003)	(0.004)	(0.006)	(0.007)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.014)	(0.011)	(0.009)	(0.009)	(0.008)
Observations	6,204	4,696	2,356	2,356	2,356	6,204	4,696	2,356	2,356	2,356	6,204	4,696	2,356	2,356	2,356
R-squared	0.003	0.004	0.006	0.010	0.012	0.030	0.038	0.070	0.086	0.090	0.011	0.016	0.025	0.029	0.035
Number of villages	1,551	1,174	589	589	589	1,551	1,174	589	589	589	1,551	1,174	589	589	589
Number of clusters	15	14	8	8	8	15	14	8	8	8	15	14	8	8	8
PANEL A: OTHER INFRASTRUCTURES															
	Public phone in village					Post office in village					Permanent market in village				
TREATMENT × 2005	0.130*	0.114*	0.112*	0.141***	0.141***	0.031***	0.035***	0.043**	0.026**	0.026**	-0.002	0.009	0.042	0.088***	0.087***
	(0.073)	(0.054)	(0.050)	(0.028)	(0.028)	(0.007)	(0.009)	(0.017)	(0.009)	(0.011)	(0.012)	(0.020)	(0.040)	(0.019)	(0.017)
TREATMENT × 2002	0.155**	0.137**	0.135**	0.198***	0.197***	-0.001	-0.001	0.005	0.003	0.002	0.013	0.012	0.017	0.007	0.006
	(0.071)	(0.063)	(0.053)	(0.028)	(0.027)	(0.004)	(0.003)	(0.004)	(0.003)	(0.002)	(0.030)	(0.030)	(0.035)	(0.026)	(0.027)
TREATMENT × 1999	0.065	0.049	0.047	0.063*	0.063*	0.022**	0.016***	0.006	0.000	0.000	0.025	0.032	0.061*	0.064***	0.064***
	(0.052)	(0.037)	(0.028)	(0.029)	(0.028)	(0.008)	(0.002)	(0.009)	(0.011)	(0.010)	(0.027)	(0.022)	(0.027)	(0.015)	(0.016)
Observations	6,204	4,696	2,356	2,356	2,356	6,204	4,696	2,356	2,356	2,356	6,204	4,696	2,356	2,356	2,356
R-squared	0.094	0.095	0.109	0.129	0.130	0.004	0.003	0.005	0.010	0.012	0.006	0.005	0.012	0.019	0.021
Number of villages	1,551	1,174	589	589	589	1,551	1,174	589	589	589	1,551	1,174	589	589	589
Number of clusters	15	14	8	8	8	15	14	8	8	8	15	14	8	8	8
Village FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE or Segment-year FE	year	year	year	segment	segment	year	year	year	segment	segment	year	year	year	segment	segment
Sample	<100km	<75km	<50km	<50km	<50km	<100km	<75km	<50km	<50km	<50km	<100km	<75km	<50km	<50km	<50km
Geographical controls					cubic					cubic					cubic

Note: robust standard errors in parentheses clustered at the district level as in 1993. *** p<0.01, ** p<0.05, * p<0.1. Columns (1)-(3), (6)-(8), (11)-(13) show the coefficient estimates associated with specification (1), i.e., the Difference-in-Difference strategy. Columns (4)-(5), (9)-(10) and (14)-(15) show the coefficient estimates associated with specification (2), i.e., the spatial Regression Discontinuity strategy.

TABLE 7

IMPACT OF RESOURCE TRANSFERS CONTROLLING FOR PRODUCING DISTRICTS: KALIMANTAN AREA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
PANEL A: EDUCATION INFRASTRUCTURES															
	Primary school in village					Junior high school					Senior high school				
TREATMENT × 2005	0.084**	0.092*	0.115*	0.089	0.088	0.125***	0.121***	0.121***	0.100***	0.099***	0.130***	0.106***	0.117***	0.109**	0.107**
	(0.037)	(0.044)	(0.050)	(0.055)	(0.057)	(0.017)	(0.034)	(0.032)	(0.024)	(0.026)	(0.012)	(0.022)	(0.033)	(0.035)	(0.039)
TREATMENT × 2002	0.017	0.002	-0.002	-0.021	-0.021	0.057***	0.043	0.044*	0.031**	0.032**	0.026	0.014	0.009	0.022*	0.021**
	(0.015)	(0.014)	(0.013)	(0.012)	(0.014)	(0.009)	(0.025)	(0.020)	(0.013)	(0.013)	(0.023)	(0.013)	(0.016)	(0.010)	(0.009)
TREATMENT × 1999	0.033*	0.035*	0.029	0.025	0.025	0.039	0.044	0.043	0.012	0.012	0.018	0.007	0.001	0.003	0.003
	(0.018)	(0.020)	(0.018)	(0.021)	(0.021)	(0.034)	(0.049)	(0.049)	(0.042)	(0.044)	(0.018)	(0.004)	(0.008)	(0.011)	(0.012)
R-squared	0.017	0.024	0.035	0.046	0.049	0.037	0.037	0.052	0.059	0.061	0.046	0.042	0.064	0.077	0.087
PANEL B: HEALTH AND OTHER INFRASTRUCTURES															
	Majority traffic through land					Health center in village					Post office in village				
TREATMENT × 2005	0.244***	0.263***	0.230**	0.227**	0.225**	0.075*	0.094*	0.116**	0.090**	0.089*	0.028***	0.031***	0.037*	0.023*	0.023
	(0.037)	(0.049)	(0.091)	(0.088)	(0.082)	(0.038)	(0.048)	(0.044)	(0.038)	(0.040)	(0.007)	(0.008)	(0.017)	(0.010)	(0.012)
TREATMENT × 2002	0.115***	0.144***	0.113	0.147**	0.147**	-0.016**	-0.008**	-0.006	-0.003	-0.004	-0.002	-0.003	0.003	0.002	0.001
	(0.036)	(0.036)	(0.073)	(0.056)	(0.054)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)	(0.005)	(0.003)	(0.003)
TREATMENT × 1999	-0.063	-0.064	-0.071	-0.012	-0.012	-0.019***	-0.010***	-0.009	-0.005	-0.005	0.021**	0.015***	0.001	-0.002	-0.002
	(0.074)	(0.076)	(0.085)	(0.072)	(0.072)	(0.004)	(0.003)	(0.005)	(0.004)	(0.004)	(0.008)	(0.002)	(0.010)	(0.010)	(0.009)
R-squared	0.113	0.131	0.141	0.156	0.166	0.031	0.039	0.072	0.088	0.092	0.004	0.004	0.006	0.011	0.013
Village FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE or Segment-year FE	year	year	year	segment	segment	year	year	year	segment	segment	year	year	year	segment	segment
Sample	<100km	<75km	<50km	<50km	<50km	<100km	<75km	<50km	<50km	<50km	<100km	<75km	<50km	<50km	<50km
Geographical controls					cubic					cubic					cubic
Observations	6,204	4,696	2,356	2,356	2,356	6,204	4,696	2,356	2,356	2,356	6,204	4,696	2,356	2,356	2,356
Number of villages	1,551	1,174	589	589	589	1,551	1,174	589	589	589	1,551	1,174	589	589	589
Number of clusters	15	14	8	8	8	15	14	8	8	8	15	14	8	8	8

Note: robust standard errors in parentheses clustered at the district level as in 1993. *** p<0.01, ** p<0.05, * p<0.1. Columns (1)-(3), (6)-(8), (11)-(13) show the coefficient estimates associated with specification (1), i.e., the Difference-in-Difference strategy. Columns (4)-(5), (9)-(10) and (14)-(15) show the coefficient estimates associated with specification (2), i.e., the spatial Regression Discontinuity strategy.

Previous doctoral theses in the Department of Economics, Gothenburg

Avhandlingar publicerade innan serien Ekonomiska Studier startades
(Theses published before the series Ekonomiska Studier was started):

Östman, Hugo (1911), Norrlands ekonomiska utveckling
Moritz, Marcus (1911), Den svenska tobaksindustrien
Sundbom, I. (1933), Prisbildning och ändamålsenlighet
Gerhard, I. (1948), Problem rörande Sveriges utrikeshandel 1936/38
Hegeland, Hugo (1951), The Quantity Theory of Money
Mattsson, Bengt (1970), Cost-Benefit analys
Rosengren, Björn (1975), Valutareglering och nationell ekonomisk politik
Hjalmarsson, Lennart (1975), Studies in a Dynamic Theory of Production and its Applications
Örtendahl, Per-Anders (1975), Substitutionsaspekter på produktionsprocessen vid massaframställning
Anderson, Arne M. (1976), Produktion, kapacitet och kostnader vid ett helautomatiskt emballageglasbruk
Ohlsson, Olle (1976), Substitution och odelbarheter i produktionsprocessen vid massaframställning
Gunnarsson, Jan (1976), Produktionssystem och tätortshierarki – om sambandet mellan rumslig och ekonomisk struktur
Köstner, Evert (1976), Optimal allokering av tid mellan utbildning och arbete
Wigren, Rune (1976), Analys av regionala effektivitetsskillnader inom industribranscher
Wästlund, Jan (1976), Skattning och analys av regionala effektivitetsskillnader inom industribranscher
Flöjstad, Gunnar (1976), Studies in Distortions, Trade and Allocation Problems
Sandelin, Bo (1977), Prisutveckling och kapitalvinster på bostadsfastigheter
Dahlberg, Lars (1977), Empirical Studies in Public Planning
Lönnroth, Johan (1977), Marxism som matematisk ekonomi
Johansson, Börje (1978), Contributions to Sequential Analysis of Oligopolistic Competition

Ekonomiska Studier, utgivna av Nationalekonomiska institutionen vid Göteborgs Universitet. Nr 1 och 4 var inte doktorsavhandlingar. (The contributions to the department series 'Ekonomiska Studier' where no. 1 and 4 were no doctoral theses):

2. **Ambjörn, Erik** (1959), Svenskt importberoende 1926-1956: en ekonomisk-statistisk kartläggning med kommentarer
3. **Landgren, K-G.** (1960), Den "Nya ekonomien" i Sverige: J.M. Keynes, E. Wigfors och utecklingen 1927-39
5. **Bigsten, Arne** (1979), Regional Inequality and Development: A Case Study of Kenya
6. **Andersson, Lars** (1979), Statens styrning av de kommunala budgetarnas struktur (Central Government Influence on the Structure of the Municipal Budget)
7. **Gustafsson, Björn** (1979), Inkomst- och uppväxtförhållanden (Income and Family Background)

8. **Granholm, Arne** (1981), Interregional Planning Models for the Allocation of Private and Public Investments
9. **Lundborg, Per** (1982), Trade Policy and Development: Income Distributional Effects in the Less Developed Countries of the US and EEC Policies for Agricultural Commodities
10. **Juås, Birgitta** (1982), Värdering av risken för personskador. En jämförande studie av implicita och explicita värden. (Valuation of Personal Injuries. A comparison of Explicit and Implicit Values)
11. **Bergendahl, Per-Anders** (1982), Energi och ekonomi - tillämpningar av input-output analys (Energy and the Economy - Applications of Input-Output Analysis)
12. **Blomström, Magnus** (1983), Foreign Investment, Technical Efficiency and Structural Change - Evidence from the Mexican Manufacturing Industry
13. **Larsson, Lars-Göran** (1983), Comparative Statics on the Basis of Optimization Methods
14. **Persson, Håkan** (1983), Theory and Applications of Multisectoral Growth Models
15. **Sterner, Thomas** (1986), Energy Use in Mexican Industry.
16. **Flood, Lennart** (1986), On the Application of Time Use and Expenditure Allocation Models.
17. **Schuller, Bernd-Joachim** (1986), Ekonomi och kriminalitet - en empirisk undersökning av brottsligheten i Sverige (Economics of crime - an empirical analysis of crime in Sweden)
18. **Walfridson, Bo** (1987), Dynamic Models of Factor Demand. An Application to Swedish Industry.
19. **Stålhammar, Nils-Olov** (1987), Strukturomvandling, företagsbeteende och förväntningsbildning inom den svenska tillverkningsindustrin (Structural Change, Firm Behaviour and Expectation Formation in Swedish Manufactury)
20. **Anxo, Dominique** (1988), Sysselsättningseffekter av en allmän arbetstidsför-kortning (Employment effects of a general shortage of the working time)
21. **Mbelle, Ammon** (1988), Foreign Exchange and Industrial Development: A Study of Tanzania.
22. **Ongaro, Wilfred** (1988), Adoption of New Farming Technology: A Case Study of Maize Production in Western Kenya.
23. **Zejan, Mario** (1988), Studies in the Behavior of Swedish Multinationals.
24. **Görling, Anders** (1988), Ekonomisk tillväxt och miljö. Förorenings-struktur och ekonomiska effekter av olika miljövårdsprogram. (Economic Growth and Environment. Pollution Structure and Economic Effects of Some Environmental Programs).
25. **Aguilar, Renato** (1988), Efficiency in Production: Theory and an Application on Kenyan Smallholders.
26. **Kayizzi-Mugerwa, Steve** (1988), External Shocks and Adjustment in Zambia.

27. **Bornmalm-Jardelöw, Gunilla** (1988), Högre utbildning och arbetsmarknad (Higher Education and the Labour Market)
28. **Tansini, Ruben** (1989), Technology Transfer: Dairy Industries in Sweden and Uruguay.
29. **Andersson, Irene** (1989), Familjebeskattning, konsumtion och arbetsutbud - En ekonometrisk analys av löne- och inkomstelasticiteter samt policysimuleringar för svenska hushåll (Family Taxation, Consumption and Labour Supply - An Econometric Analysis of Wage and Income Elasticities and Policy Simulations for Swedish Households)
30. **Henrekson, Magnus** (1990), An Economic Analysis of Swedish Government Expenditure
31. **Sjöö, Boo** (1990), Monetary Policy in a Continuous Time Dynamic Model for Sweden
32. **Rosén, Åsa** (1991), Contributions to the Theory of Labour Contracts.
33. **Loureiro, Joao M. de Matos** (1992), Foreign Exchange Intervention, Sterilization and Credibility in the EMS: An Empirical Study
34. **Irandoust, Manuchehr** (1993), Essays on the Behavior and Performance of the Car Industry
35. **Tasiran, Ali Cevat** (1993), Wage and Income Effects on the Timing and Spacing of Births in Sweden and the United States
36. **Milopoulos, Christos** (1993), Investment Behaviour under Uncertainty: An Econometric Analysis of Swedish Panel Data
37. **Andersson, Per-Åke** (1993), Labour Market Structure in a Controlled Economy: The Case of Zambia
38. **Storrie, Donald W.** (1993), The Anatomy of a Large Swedish Plant Closure
39. **Semboja, Haji Hatibu Haji** (1993), Energy and Development in Kenya
40. **Makonnen, Negatu** (1993), Labor Supply and the Distribution of Economic Well-Being: A Case Study of Lesotho
41. **Julin, Eva** (1993), Structural Change in Rural Kenya
42. **Durevall, Dick** (1993), Essays on Chronic Inflation: The Brazilian Experience
43. **Veiderpass, Ann** (1993), Swedish Retail Electricity Distribution: A Non-Parametric Approach to Efficiency and Productivity Change
44. **Odeck, James** (1993), Measuring Productivity Growth and Efficiency with Data Envelopment Analysis: An Application on the Norwegian Road Sector
45. **Mwenda, Abraham** (1993), Credit Rationing and Investment Behaviour under Market Imperfections: Evidence from Commercial Agriculture in Zambia
46. **Mlambo, Kupukile** (1993), Total Factor Productivity Growth: An Empirical Analysis of Zimbabwe's Manufacturing Sector Based on Factor Demand Modelling
47. **Ndung'u, Njuguna** (1993), Dynamics of the Inflationary Process in Kenya
48. **Modén, Karl-Markus** (1993), Tax Incentives of Corporate Mergers and Foreign Direct Investments
49. **Franzén, Mikael** (1994), Gasoline Demand - A Comparison of Models
50. **Heshmati, Almas** (1994), Estimating Technical Efficiency, Productivity Growth And Selectivity Bias Using Rotating Panel Data: An Application to Swedish Agriculture

51. **Salas, Osvaldo** (1994), Efficiency and Productivity Change: A Micro Data Case Study of the Colombian Cement Industry
52. **Bjurek, Hans** (1994), Essays on Efficiency and Productivity Change with Applications to Public Service Production
53. **Cabezas Vega, Luis** (1994), Factor Substitution, Capacity Utilization and Total Factor Productivity Growth in the Peruvian Manufacturing Industry
54. **Katz, Katarina** (1994), Gender Differentiation and Discrimination. A Study of Soviet Wages
55. **Asal, Maher** (1995), Real Exchange Rate Determination and the Adjustment Process: An Empirical Study in the Cases of Sweden and Egypt
56. **Kjulin, Urban** (1995), Economic Perspectives on Child Care
57. **Andersson, Göran** (1995), Volatility Forecasting and Efficiency of the Swedish Call Options Market
58. **Forteza, Alvaro** (1996), Credibility, Inflation and Incentive Distortions in the Welfare State
59. **Locking, Håkan** (1996), Essays on Swedish Wage Formation
60. **Välilä, Timo** (1996), Essays on the Credibility of Central Bank Independence
61. **Yilma, Mulugeta** (1996), Measuring Smallholder Efficiency: Ugandan Coffee and Food-Crop Production
62. **Mabugu, Ramos E.** (1996), Tax Policy Analysis in Zimbabwe Applying General Equilibrium Models
63. **Johansson, Olof** (1996), Welfare, Externalities, and Taxation; Theory and Some Road Transport Applications.
64. **Chitiga, Margaret** (1996), Computable General Equilibrium Analysis of Income Distribution Policies in Zimbabwe
65. **Leander, Per** (1996), Foreign Exchange Market Behavior Expectations and Chaos
66. **Hansen, Jörgen** (1997), Essays on Earnings and Labor Supply
67. **Cotfas, Mihai** (1997), Essays on Productivity and Efficiency in the Romanian Cement Industry
68. **Horgby, Per-Johan** (1997), Essays on Sharing, Management and Evaluation of Health Risks
69. **Nafar, Nosratollah** (1997), Efficiency and Productivity in Iranian Manufacturing Industries
70. **Zheng, Jinghai** (1997), Essays on Industrial Structure, Technical Change, Employment Adjustment, and Technical Efficiency
71. **Isaksson, Anders** (1997), Essays on Financial Liberalisation in Developing Countries: Capital mobility, price stability, and savings
72. **Gerdin, Anders** (1997), On Productivity and Growth in Kenya, 1964-94
73. **Sharifi, Alimorad** (1998), The Electricity Supply Industry in Iran: Organization, performance and future development
74. **Zamanian, Max** (1997), Methods for Mutual Fund Portfolio Evaluation: An application to the Swedish market
75. **Manda, Damiano Kulundu** (1997), Labour Supply, Returns to Education, and the Effect of Firm Size on Wages: The case of Kenya
76. **Holmén, Martin** (1998), Essays on Corporate Acquisitions and Stock Market Introductions

77. **Pan, Kelvin** (1998), Essays on Enforcement in Money and Banking
78. **Rogat, Jorge** (1998), The Value of Improved Air Quality in Santiago de Chile
79. **Peterson, Stefan** (1998), Essays on Large Shareholders and Corporate Control
80. **Belhaj, Mohammed** (1998), Energy, Transportation and Urban Environment in Africa: The Case of Rabat-Salé, Morocco
81. **Mekonnen, Alemu** (1998), Rural Energy and Afforestation: Case Studies from Ethiopia
82. **Johansson, Anders** (1998), Empirical Essays on Financial and Real Investment Behavior
83. **Köhlin, Gunnar** (1998), The Value of Social Forestry in Orissa, India
84. **Levin, Jörgen** (1998), Structural Adjustment and Poverty: The Case of Kenya
85. **Ncube, Mkhululi** (1998), Analysis of Employment Behaviour in Zimbabwe
86. **Mwansa, Ladslous** (1998), Determinants of Inflation in Zambia
87. **Agnarsson, Sveinn** (1998), Of Men and Machines: Essays in Applied Labour and Production Economics
88. **Kadenge, Phineas** (1998), Essays on Macroeconomic Adjustment in Zimbabwe: Inflation, Money Demand, and the Real Exchange Rate
89. **Nyman, Håkan** (1998), An Economic Analysis of Lone Motherhood in Sweden
90. **Carlsson, Fredrik** (1999), Essays on Externalities and Transport
91. **Johansson, Mats** (1999), Empirical Studies of Income Distribution
92. **Alemu, Tekie** (1999), Land Tenure and Soil Conservation: Evidence from Ethiopia
93. **Lundvall, Karl** (1999), Essays on Manufacturing Production in a Developing Economy: Kenya 1992-94
94. **Zhang, Jianhua** (1999), Essays on Emerging Market Finance
95. **Mlima, Aziz Ponary** (1999), Four Essays on Efficiency and Productivity in Swedish Banking
96. **Davidsen, Björn-Ivar** (2000), Bidrag til den økonomisk-metodologiske tenkningen (Contributions to the Economic Methodological Thinking)
97. **Ericson, Peter** (2000), Essays on Labor Supply
98. **Söderbom, Måns** (2000), Investment in African Manufacturing: A Microeconomic Analysis
99. **Höglund, Lena** (2000), Essays on Environmental Regulation with Applications to Sweden
100. **Olsson, Ola** (2000), Perspectives on Knowledge and Growth
101. **Meuller, Lars** (2000), Essays on Money and Credit
102. **Österberg, Torun** (2000), Economic Perspectives on Immigrants and Intergenerational Transmissions
103. **Kalinda Mkenda, Beatrice** (2001), Essays on Purchasing Power Parity, RealExchange Rate, and Optimum Currency Areas
104. **Nerhagen, Lena** (2001), Travel Demand and Value of Time - Towards an Understanding of Individuals Choice Behavior

105. **Mkenda, Adolf** (2001), Fishery Resources and Welfare in Rural Zanzibar
106. **Eggert, Håkan** (2001), Essays on Fisheries Economics
107. **Andrén, Daniela** (2001), Work, Sickness, Earnings, and Early Exits from the Labor Market. An Empirical Analysis Using Swedish Longitudinal Data
108. **Nivorozhkin, Eugene** (2001), Essays on Capital Structure
109. **Hammar, Henrik** (2001), Essays on Policy Instruments: Applications to Smoking and the Environment
110. **Nannyonjo, Justine** (2002), Financial Sector Reforms in Uganda (1990-2000): Interest Rate Spreads, Market Structure, Bank Performance and Monetary Policy
111. **Wu, Hong** (2002), Essays on Insurance Economics
112. **Linde-Rahr, Martin** (2002), Household Economics of Agriculture and Forestry in Rural Vienam
113. **Maneschiöld, Per-Ola** (2002), Essays on Exchange Rates and Central Bank Credibility
114. **Andrén, Thomas** (2002), Essays on Training, Welfare and Labor Supply
115. **Granér, Mats** (2002), Essays on Trade and Productivity: Case Studies of Manufacturing in Chile and Kenya
116. **Jaldell, Henrik** (2002), Essays on the Performance of Fire and Rescue Services
117. **Alpizar, Francisco, R.** (2002), Essays on Environmental Policy-Making in Developing Countries: Applications to Costa Rica
118. **Wahlberg, Roger** (2002), Essays on Discrimination, Welfare and Labor Supply
119. **Piculescu, Violeta** (2002), Studies on the Post-Communist Transition
120. **Pylkkänen, Elina** (2003), Studies on Household Labor Supply and Home Production
121. **Löfgren, Åsa** (2003), Environmental Taxation – Empirical and Theoretical Applications
122. **Ivaschenko, Oleksiy** (2003), Essays on Poverty, Income Inequality and Health in Transition Economies
123. **Lundström, Susanna** (2003), On Institutions, Economic Growth and the Environment
124. **Wambugu, Anthony** (2003), Essays on Earnings and Human Capital in Kenya
125. **Adler, Johan** (2003), Aspects of Macroeconomic Saving
126. **Erlandsson, Mattias** (2003), On Monetary Integration and Macroeconomic Policy
127. **Brink, Anna** (2003), On the Political Economy of Municipality Break-Ups
128. **Ljungwall, Christer** (2003), Essays on China's Economic Performance During the Reform Period
129. **Chifamba, Ronald** (2003), Analysis of Mining Investments in Zimbabwe
130. **Muchapondwa, Edwin** (2003), The Economics of Community-Based Wildlife Conservation in Zimbabwe
131. **Hammes, Klaus** (2003), Essays on Capital Structure and Trade Financing
132. **Abou-Ali, Hala** (2003), Water and Health in Egypt: An Empirical Analysis
133. **Simatele, Munacinga** (2004), Financial Sector Reforms and Monetary Policy in Zambia
134. **Tezic, Kerem** (2004), Essays on Immigrants' Economic Integration
135. INSTÄLLD

136. **Gjirja, Matilda** (2004), Efficiency and Productivity in Swedish Banking
137. **Andersson, Jessica** (2004), Welfare Environment and Tourism in Developing Countries
138. **Chen, Yinghong** (2004), Essays on Voting Power, Corporate Governance and Capital Structure
139. **Yesuf, Mahmud** (2004), Risk, Time and Land Management under Market Imperfections: Applications to Ethiopia
140. **Kateregga, Eseza** (2005), Essays on the Infestation of Lake Victoria by the Water Hyacinth
141. **Edvardsen, Dag Fjeld** (2004), Four Essays on the Measurement of Productive Efficiency
142. **Lidén, Erik** (2005), Essays on Information and Conflicts of Interest in Stock Recommendations
143. **Dieden, Sten** (2005), Income Generation in the African and Coloured Population – Three Essays on the Origins of Household Incomes in South Africa
144. **Eliasson, Marcus** (2005), Individual and Family Consequences of Involuntary Job Loss
145. **Mahmud, Minhaj** (2005), Measuring Trust and the Value of Statistical Lives: Evidence from Bangladesh
146. **Lokina, Razack Bakari** (2005), Efficiency, Risk and Regulation Compliance: Applications to Lake Victoria Fisheries in Tanzania
147. **Jussila Hammes, Johanna** (2005), Essays on the Political Economy of Land Use Change
148. **Nyangena, Wilfred** (2006), Essays on Soil Conservation, Social Capital and Technology Adoption
149. **Nivorozhkin, Anton** (2006), Essays on Unemployment Duration and Programme Evaluation
150. **Sandén, Klas** (2006), Essays on the Skill Premium
151. **Deng, Daniel** (2006), Three Essays on Electricity Spot and Financial Derivative Prices at the Nordic Power Exchange
152. **Gebreeyesus, Mulu** (2006), Essays on Firm Turnover, Growth, and Investment Behavior in Ethiopian Manufacturing
153. **Islam, Nizamul Md.** (2006), Essays on Labor Supply and Poverty: A Microeconometric Application
154. **Kjaer, Mats** (2006), Pricing of Some Path-Dependent Options on Equities and Commodities
155. **Shimeles, Abebe** (2006), Essays on Poverty, Risk and Consumption Dynamics in Ethiopia
156. **Larsson, Jan** (2006), Four Essays on Technology, Productivity and Environment
157. **Congdon Fors, Heather** (2006), Essays in Institutional and Development Economics
158. **Akpalu, Wisdom** (2006), Essays on Economics of Natural Resource Management and Experiments
159. **Daruvala, Dinky** (2006), Experimental Studies on Risk, Inequality and Relative Standing
160. **García, Jorge** (2007), Essays on Asymmetric Information and Environmental Regulation through Disclosure

161. **Bezabih, Mintewab** (2007), Essays on Land Lease Markets, Productivity, Biodiversity, and Environmental Variability

162. **Visser, Martine** (2007), Fairness, Reciprocity and Inequality: Experimental Evidence from South Africa

163. **Holm, Louise** (2007), A Non-Stationary Perspective on the European and Swedish Business Cycle

164. **Herbertsson, Alexander** (2007), Pricing Portfolio Credit Derivatives

165. **Johansson, Anders C.** (2007), Essays in Empirical Finance: Volatility, Interdependencies, and Risk in Emerging Markets

166. **Ibáñez Díaz, Marcela** (2007), Social Dilemmas: The Role of Incentives, Norms and Institutions

167. **Ekbom, Anders** (2007), Economic Analysis of Soil Capital, Land Use and Agricultural Production in Kenya

168. **Sjöberg, Pål** (2007), Essays on Performance and Growth in Swedish Banking

169. **Palma Aguirre, Grisha Alexis** (2008), Explaining Earnings and Income Inequality in Chile

170. **Akay, Alpaslan** (2008), Essays on Microeconometrics and Immigrant Assimilation

171. **Carlsson, Evert** (2008), After Work – Investing for Retirement

172. **Munshi, Farzana** (2008), Essays on Globalization and Occupational Wages

173. **Tsakas, Elias** (2008), Essays on Epistemology and Evolutionary Game Theory

174. **Erlandzon, Karl** (2008), Retirement Planning: Portfolio Choice for Long-Term Investors

175. **Lampi, Elina** (2008), Individual Preferences, Choices, and Risk Perceptions – Survey Based Evidence

176. **Mitrut, Andreea** (2008), Four Essays on Interhousehold Transfers and Institutions in Post-Communist Romania

177. **Hansson, Gustav** (2008), Essays on Social Distance, Institutions, and Economic Growth

178. **Zikhali, Precious** (2008), Land Reform, Trust and Natural Resource Management in Africa

179. **Tengstam, Sven** (2008), Essays on Smallholder Diversification, Industry Location, Debt Relief, and Disability and Utility

180. **Boman, Anders** (2009), Geographic Labour Mobility – Causes and Consequences

181. **Qin, Ping** (2009), Risk, Relative Standing and Property Rights: Rural Household Decision-Making in China

182. **Wei, Jiegen** (2009), Essays in Climate Change and Forest Management

183. **Belu, Constantin** (2009), Essays on Efficiency Measurement and Corporate Social Responsibility

184. **Ahlerup, Pelle** (2009), Essays on Conflict, Institutions, and Ethnic Diversity

185. **Quiroga, Miguel** (2009), Microeconomic Policy for Development: Essays on Trade and Environment, Poverty and Education

186. **Zerfu, Daniel** (2010), Essays on Institutions and Economic Outcomes

187. **Wollbrant, Conny** (2010), Self-Control and Altruism

188. **Villegas Palacio, Clara** (2010), Formal and Informal Regulations: Enforcement and Compliance

189. **Maican, Florin** (2010), Essays in Industry Dynamics on Imperfectly Competitive

Markets

190. **Jakobsson, Niklas** (2010), Laws, Attitudes and Public Policy
191. **Manescu, Cristiana** (2010), Economic Implications of Corporate Social Responsibility and Responsible Investments
192. **He, Haoran** (2010), Environmental and Behavioral Economics – Applications to China
193. **Andersson, Fredrik W.** (2011), Essays on Social Comparison
194. **Isaksson, Ann-Sofie** (2011), Essays on Institutions, Inequality and Development
195. **Pham, Khanh Nam** (2011), Prosocial Behavior, Social Interaction and Development: Experimental Evidence from Vietnam
196. **Lindskog, Annika** (2011), Essays on Economic Behaviour: HIV/AIDS, Schooling, and Inequality
197. **Kotsadam, Andreas** (2011), Gender, Work, and Attitudes
198. **Alem, Yonas** (2011), Essays on Shocks, Welfare, and Poverty Dynamics: Microeconometric Evidence from Ethiopia
199. **Köksal-Ayhan, Miyase Yesim** (2011), Parallel Trade, Reference Pricing and Competition in the Pharmaceutical Market: Theory and Evidence
200. **Vondolia, Godwin Kofi** (2011), Essays on Natural Resource Economics
201. **Widerberg, Anna** (2011), Essays on Energy and Climate Policy – Green Certificates, Emissions Trading and Electricity Prices
202. **Siba, Eyerusalem** (2011), Essays on Industrial Development and Political Economy of Africa
203. **Orth, Matilda** (2012), Entry, Competition and Productivity in Retail
204. **Nerman, Måns** (2012), Essays on Development: Household Income, Education, and Female Participation and Representation
205. **Wicks, Rick** (2012), The Place of Conventional Economics in a World with Communities and Social Goods
206. **Sato, Yoshihiro** (2012), Dynamic Investment Models, Employment Generation and Productivity – Evidence from Swedish Data
207. **Valsecchi, Michele** (2012), Essays in Political Economy of Development