The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

Conservation Contracts and Political Regimes

Bård Harstad
Torben Mideksa

CESifo Working Paper No. 5334
Category 9: Resource and Environment Economics
April 2015

An electronic version of the paper may be downloaded
- *from the SSRN website: www.SSRN.com*
- *from the RePEc website: www.RePEc.org*
- *from the CESifo website: www.CESifo-group.org/wp*

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 5334*

Conservation Contracts and Political Regimes

Abstract

Motivated by tropical deforestation, we analyze (i) a novel theory of resource extraction, (ii) the optimal conservation contract, (iii) when the donor prefers contracting with central rather than local governments, and (iv) how the donor's presence may induce institutional change. Deforestation can be legal or illegal in the model: each district decides how much to protect and how much to extract for sale on a common market. If districts are strong, in that they find protection inexpensive, extraction is sales-driven and districts bene.t if neighbors conserve. If districts are weak, they lose when neighbors conserve since the smaller supply increases the price and the pressure on the resource, and thus also the cost of protection. Consequently, decentralizing authority increases conservation if and only if districts are weak. Contracting with the central authority is socially optimal, but, on the one hand, the donor benefits from contracting with districts if they are weak; on the other hand, districts prefer to decentralize if they are strong. The presence of the donor may lead to a regime change that increases extraction by more than it is reduced by the contract itself.

JEL-Code: Q340.

Keywords: deforestation, resource extraction, conservation, contracts, centralization, decentralization, externalities, participation constraints, incentive constraints, tropical forests, climate change, REDD, PES.

Bård Harstad
University of Oslo / Norway
bard.harstad@econ.uio.no

Torben Mideksa
University of Oslo / Norway
torben.mideksa@econ.uio.no

29 April 2015
We are grateful to audiences at Harvard University, London School of Economics, Warwick University, University of Zurich, Toulouse School of Economics, Paris Ecole Polytechnique, Universidad Carlos III, University of Copenhagen, University of Oslo, and the 2015 Environmental Protection and Sustainability Forum in Bath. We are especially thankful to the comments of Arild Angelsen, Philippe Delacote, Jonas Hjort, Chuck Mason, Nicola Persico, Torsen Persson, Francois Salanie, Steve Shavell, and Kathryn Spier. Judith Levy assisted with the editing.

1 Introduction

Natural resources are being depleted all across the world. They are managed and extracted by independent countries, even though conservation may also benefit third parties. The economics and politics of resource extraction are inextricable and must be better understood before conservation can succeed. This paper presents a new model of conservation and derives the contract preferred by a third party who benefits from conservation. We also show how the contract both influences, and should be influenced by, the countries' political regimes and state capacities. Payments for environmental services (PES) are important in many settings, and the resource in our model could be fossil fuels or land use quite generally, but our analysis is motivated in particular by deforestation in the tropics and the emergence of contracts on reducing emissions from deforestation and forest degradation (REDD).[1]

Deforestation in the tropics is an immensely important problem. The cumulative effect of deforestation amounts to about one quarter of anthropogenic greenhouse gas emissions, which generate global warming (Edenhofer et al., 2014). The annual contribution from deforestation to CO_2 emissions is around ten percent (Stocker et al., 2013), and more when other greenhouse gases are taken into account. In addition to the effect on global warming, deforestation leads to huge losses in biodiversity. The negative externalities of deforestation amount to $2-4.5 trillion a year, according to The Economist (2010). Nevertheless, Hansen et al. (2013) document that tropical forest loss has been *increasing* at an average rate of 2101 km^2 each year since 2000.[2]

Third parties are therefore interested in conservation. With the help of donor countries (in particular Norway, Germany, and Japan), the World Bank and the United Nations are already offering financial incentives to reduce deforestation in a number of countries. Estimates suggest that deforestation could be halved at a cost of $21-35 billion per year, or reduced by 20-30 percent at a price at $10/tCO_2$.[3] Conservation contracts are thus likely to be an important part of future climate change policies and treaties. They are also favored by economists who view them as the natural Coasian solution (Alston and Andersson, 2011). Even for other types of resources, such as fossil

[1] See Engel et al. (2008) for PES more generally, and see Karsenty (2008) and Parker et al. (2009) for an explanation of the difference between alternative concepts such as RED, REDD, and REDD+.

[2] Harris et al. (2012) offer more precise estimates of deforestation between 2000 and 2005. The overall message that tropical deforestation has been increasing remains robust.

[3] See, respectively, Edenhofer et al. (2014) and Busch et al. (2012).

fuel reserves, recent theory suggests that the very best climate policy is for the climate coalition to pay nonparticipants to conserve particular reserves (Harstad, 2012).

It is therefore important to understand how conservation contracts should be designed. So far, however, there is little theory that can guide real-world contract designers. At the same time, it is clear that the causes of deforestation differ across countries and regions. While local governments sell logging concessions in some countries, other countries fight illegal logging for timber or the burning of the forest for agriculture. The political regime also seems to be important: while Burgess et al. (2012) find that decentralization led to more logging in Indonesia,[4] the reverse has been documented for other regions, such as the Himalayas.[5] Despite these differences, contracts tend to be similar across countries, and targeted at the national government only. Norway, for example, has recently declined to contract with the region Madre de Dios in Peru. Perhaps, after all, the very existence of REDD contracts motivate districts to centralize authority, as suggested by Phelps et al. (2010).

These facts and claims raise a number of essential questions. How can we explain the inconsistent effect of the political regime on conservation? What is the optimal conservation contract, and how does it depend on state capacities or the driver of deforestation? Is it wise to contract with central governments only, or can local contracts be more effective? Can the existence of conservation contracts indeed influence regime change, and when would that be beneficial and increase conservation?

Our first contribution is to provide a simple, innovative model of conservation that can address all the questions above. Each district may benefit from extracting its resource, but the price of the harvest is reduced by the districts' aggregate supply. To protect the remaining part of the resource, the monitoring effort must ensure that the expected penalty is at least as large as the harvest price motivating illegal logging. Thus, a district may want to limit the amount that is protected, and let some of it be harvested and offered to the market, since this reduces the price and thus the monitoring cost on the part that is to be protected.

The model can explain the inconsistent evidence regarding the effect of the political regime. Suppose districts are "strong" in that extraction is sales-driven and motivated

[4]Kaimovitz et al. (1998) and Ribot (2002) also report evidence of a similarly strong effect of decentralization on deforestation.

[5]See Antinori and Rausser (2007), Baland et al. (2010), Chhatre and Agrawal (2008), and Somanathan et al. (2009).

mainly by the profit that can be earned by the districts. In this case, a district benefits if the neighbors conserve since that reduces their supply and the harvest price (and profit) increases. The positive (pecuniary) externality from conservation would be internalized by a central government, so centralizing authority will increase conservation. Alternatively, suppose logging is illegal or districts are "weak" in that they are unable to capture much of the profit, and they find it expensive to protect the resource. In that case, a district loses when neighbors conserve, since this increases the price and the pressure on the resource, and thus also the monitoring cost when the resource is protected. The negative externality implies that when authority is centralized, conservation declines.

Our second contribution is to derive the optimal conservation contract. If there is a single district, a simple contract (similar to a Pigou subsidy) implements the first-best, regardless of the other parameters in the model. This finding, in isolation, supports today's use of contracts that are linear in the amount of avoided deforestation. With multiple districts, however, one district finds it optimal to extract more when the neighbors conserve or sign conservation contracts: the higher price makes it profitable to extract if the district is strong, and expensive to protect if the district is weak. The larger is the positive externality from conservation in one district on the other districts, the more each district will demand before adhering to a conservation contract. This cost makes it expensive for the donor to encourage conservation, and equilibrium conservation is smaller than at the first best. When the districts are weak and the externality negative, however, there is too much conservation in equilibrium, since the donor takes advantage of the negative externality on the other districts for each conservation contract that is offered.

When districts are strong and extraction is sales-driven, the positive externality means that a central government would be more willing to conserve and adhere to a conservation contract. This willingness can be exploited by the donor, who in this case prefers to contract with central governments rather than with local governments. This result is reversed when the externality is negative, i.e., when districts are weak and extraction is protection-driven. The negative externality means that a central government becomes reluctant to the conservation contract, and the donor will find that districts adhere to the contract at a lower price. Thus, the donor prefers local contracts if and only if the districts are weak.

The districts themselves, however, have the exact opposite interest. If they could, they would choose a political regime that forces the donor to offer larger transfers for each unit that is conserved. Consequently, the districts prefer to centralize authority when they are weak, and decentralize when they are strong. The presence of the donor will contribute to these incentives for institutional change. The regime change will not only make the conservation contracts expensive, it will also increase extraction. Our final contribution is to specify conditions under which the induced institutional change increases extraction by more than the conservation contracts reduce it. Under these conditions, the presence of the donor does more harm than good.

Our model and findings are consistent with the empirical literature cited above and below. When the levels of the parameters in our model are suitably adjusted, it can be applied whether the deforestation driver is corruption, revenue generation at the local level, illegal logging by small farmers, or by large corporations. A higher price (on timber or agricultural products) is in any case increasing deforestation (Kaimowitz and Angelsen, 1998), and conservation in one district increases the pressure and deforestation on other areas, as empirically documented.[6] Our analysis can explain the puzzle of why decentralization reduces protection-driven deforestation in places like Himalaya, while increasing it in places like Indonesia where districts are strong and deforestation is sales-driven.[7] We can also explain why the existing linear (and Pigou-like) contract appears to be a good choice at first, before leakage and multiple drivers are taken into account. Finally, we provide partial but limited support to the strategy of insisting on national contracts, and the claim that REDD contracts themselves may motivate centralization (Phelps et al., 2010).

We also contribute to the formal theory literature on the causes of resource extraction and deforestation. The literature points to optimal land use models, income growth and demand for forest products, corruption, costly enforcement, illegal logging, and other institutional weaknesses.[8] Alemagi and Kozak (2010) argue that several or

[6]The estimated numbers of such leakage vary between 5 and 95 percent, but more typical estimates are around 40 percent (Atmadja and Verchot, 2012; Jacobson, 2014; Meyfroidt and Lambin, 2009; Murray, 2008; Murray et al., 2004) .

[7]In Nepal, "the Forest Department was poorly staffed and thus unable to implement and enforce the national policies, and deforestation increased in the 1960s and 1970s" (Shyamsundar and Ghate, 2014, pp.85). In contrast, "Deforestation in Indonesia is largely driven by the expansion of profitable and legally sanctioned oil palm and timber plantations and logging operations" (Busch et al., 2015, pp. 1328). The literature on state capacity is often referring to states in East Asia as strong; see the references in Acemoglu et al. (2015).

[8]See, for optimal land use models: Hartwick et al. (2001); income growth and demand for forest

all of these factors might be at play in driving deforestation. Our first contribution to this literature is to provide a tractable workhorse model that can be used to study all these alternative drivers.

Our theoretical framework draws from, and in fact ties together the literatures on state capacity, the resource curse, and crime displacement. Our terminology strong vs. weak states (or districts) is borrowed from the literature on state capacity (Acemoglu, 2005; Besley and Persson, 2009, 2011), which refers to states as weak if they are unable to control the economy, support private markets, or raise revenues. The role of institutions has also been pointed out by the literature on the resource curse, which has found that a larger resource stock is beneficial for a country with good institutions, but not if the institutions are weak.[9] Our theory is in line with this idea, since strong districts can decide how to profit from its resource, while weak districts find it expensive to protect a large resource stock. Conservation contracts will aggravate the dependence on institutions, we show, since the contracts generate positive externalities when districts are strong, but negative externalities when then are weak.

The papers above are not considering the interaction between districts. In a recent paper, however, Acemoglu et al. (2015) present a model with multiple states to investigate spillovers across jurisdictions and whether state capacity are substitutes or complements when it comes to public good provision. State capacity is exogenously given in our model; instead we emphasize how the capacity determines the sign of the externality from resource extraction. Regardless of this sign, conservation in one district leads to deforestation in others. Estimates of such "leakage" can be quite high, as mentioned above, and scientists have pointed out the importance of accounting for leakage when comparing various types of REDD contracts (Busch et al., 2012).[10]

In our model, the leakage is related to shifts in market shares when districts are strong, and to crime displacement when districts are weak and the extraction is illegal. There is plenty of empirical support for crime displacement,[11] although the

products: Foster and Rosenzweig (2003); corruption: Burgess et al. (2012); Amacher et al. (2012); Delacote (2011); Robinson and Lokina (2011); costly enforcement: Clarke et al. (1993); Dokken et al. (2014); illegal logging: Amacher et al. (2007); Clarke et al. (1993); McAllister et al. (2000); Robinson and Lokina (2012); Robinson et al. (2013); or for other institutional weaknesses: Angelsen (2001); Mendelsohn (1994). Kaimowitz and Angelsen (1998) and Angelsen and Kaimowitz (1999) provide a detailed review of the earlier literature regarding economic models of tropical deforestation.

[9] See Mehlum et al. (2006), Torvik (2009), Robinson et al. (2006), Robinson et al. (2014), or the survey by van der Ploeg (2011).

[10] In contrast to the estimates in that paper, we analytically derive the optimal contract in a setting which also allows for illegal logging and protection costs.

[11] See Gonzalez-Navarro (2013) for recent evidence or the handbook chapters by Helsley (2004),

mechanism is not always through the market, as emphasized here. To the best of our knowledge, our model is the first to recognize that by letting a fraction of the resource be unprotected, the supply (of the harvest) increases and the price declines, thereby reducing the pressure and the enforcement cost on the part that is to be conserved. This mechanism may add a new perspective also to the more general literatures on crime, enforcement, and inspection games.[12]

Our second contribution, on contracts, is adding to a different literature. The leakage and the associated externality between the districts weaken the effectiveness, and influence the design of conservation contracts. We thus diverge from the growing literature on how to design agreements for PES (Engel et al., 2008) or REDD (Kerr, 2013), which tends to focus on textbook contract-theoretic problems such as moral hazard (Gjertsen et al., 2010), private information (Chiroleu-Assouline et al., 2012; Mason, 2013; Mason and Plantinga, 2013), or observability (Delacote and Simonet, 2013).[13] Instead, the analysis in our paper is more related to the literature on contracts in the presence of externalities. While the general theory has been outlined by Segal (1999), our model endogenizes the sign and the level of the externality and our results are more detailed in characterizing the contract for the particular case of resource extraction.[14] More importantly, we go further than Segal (1999) by, first, searching for the principal's optimal contracting partner (central vs. local governments), which is an important issue for real-world conservation contracts,[15] and, second, by showing how the principal's presence influences the organizational structure among the agents

Epple and Nechyba (2004), or Johnson et al. (2012) for surveys.

[12]For overviews, see Polinsky and Shavell (2007) and Avenhaus et al. (2002). Also Eeckhout et al. (2010) and Lando and Shavell (2004) reach the conclusion that it may be optimal to monitor intensively some places (or groups), and not at all elsewhere. The reason is, as in this paper, that enforcement must reach a certain level to have any impact. However, these papers do not take into account that by abstaining from monitoring some places, the required monitoring level declines for the places where the law is to be enforced. This effect, which we emphasize, means that there is an interior solution for the amount of area that is to be protected even when there is no budget constraint or convex effort cost. Our mechanism also differs from that in Kremer and Morcom (2000), where the regulator may want to increase the (potential) supply–not to reduce the monitoring cost, as here–but in order to reduce the incentive to poach and thus eliminate the bad equilibrium in a dynamic game with multiple equilibria (one of them being extinction and thus low supply).

[13]Harstad (2015) takes a political-economy (and game-theoretic) approach by showing when and why conservation contracts are not offered in equilibrium in a dynamic setting. That mechanism does not appear in the present framework, however.

[14]To be specific, we derive conditions under which the donor does not lose from using linear REDD contracts, of the type currently in place. For linear contracts, we show how the optimal reference level should generally differ from the business-as-usual level, in contrast to the traditional presumption and advise (Busch et al., 2012).

[15]See the review by Angelsen (2008), although the arguments therein do not include the effect on the rents obtained by the districts, which is the driving force behind our results.

(i.e., the districts).[16]

Our endogenization of the political regime is thus the third, and most interesting contribution of our paper. In public economics, the benefit from centralizing political power in the presence of externalities has been recognized at least since the famous decentralization theorem of Oates (1972), but our approach points to trade-offs that are better discussed in the literature on mergers in industrial organization. (After all, Cournot competition is a special case of our model.) It is well known from this literature that mergers and acquisitions can be beneficial for firms that seek monopoly power, although the actions of fringe firms must also be taken into account. In an industry of n firms, Salant et al. (1983) showed that it is not profitable for two firms to merge, since the other $n-2$ firms will produce more, as a result.[17] A merger is profitable only when it includes more than $n/2$ firms, according to Gaudet and Salant (1991). These results arise as a special case in our model, but we permit the entities to be weak rather than strong, and we consider a regulatory agency (or donor) that affects the decision to centralize. Our analysis is thus drawing on several strands of literatures to better understand the economics and the politics of conservation.

The next section presents our model of conservation and extraction before we solve it in Section 3. Conservation contracts are analyzed in Section 4, while Section 5 endogenizes the political regime and studies when the donor prefers contracting with central rather than local governments. After a brief concluding section, the Appendix provides all the proofs.

2 A Theory of Conservation

This section presents a framework with conservation and resource extraction in which there are many districts and a common market for the harvest. The framework is general in that the resource can be any kind of resource (for example, oil or land), the harvest can be timber or agricultural products, and the districts can be either countries or regions. To fix ideas, however, we refer to the resource as forest.

To motivate the framework we start by sequentially presenting two alternative models of conservation before we combine them. In both cases we have $n \geq 1$ districts

[16] Genicot and Ray (2006) allow agents to coordinate (but not centralize), and show that the principal still manages to "split and rule."

[17] Perry and Porter (1985) and Kamien and Zang (1990) change the assumptions of Salant et al. and get different results.

and we let x_i be the extraction level in district $i \in N = \{1, ..., n\}$. The aggregate harvest, $x = \sum_{i \in N} x_i$, is sold on a common market.

A sales-driven model. If districts are motivated by the profit generated by the sales, district i's payoff may be represented by $bpx_i - v_i x_i$, where b is the benefit of profit, p is the price for the harvest, and v_i is district i's marginal opportunity value when losing the forest. For example, v_i may represent the environmental benefits which the forest provides to i, the expected (and discounted) future harvest (or forest) price, or the tax or lost transfer which i experiences from more extraction. In the next section, we will let $v_i \equiv v + t_i$, where t_i is a tax or a transfer.

In this model, districts extract to raise revenues, taking into account that the more they extract, the smaller is the price. In the simplest possible setting, a linear demand curve can be derived from quadratic utility functions:

$$p = \overline{p} - ax, \tag{1}$$

where $\overline{p}$ and a are positive constants.

A protection-driven model. Consider now a setting where districts do not extract to sell, but where they try to prevent illegal extraction. If protection is difficult, one must take into account that an illegal logger earns the price p by extracting a unit of the forest. This profit must be compared to the expected penalty, θ, which one faces when logging on that unit of the forest. The enforcement is preventive if and only if the expected penalty is larger than the benefit:

$$\theta \geq p. \tag{2}$$

We let districts set their expected penalties in advance in order to discourage extraction. In principle, the expected penalty can be increased by a larger fine or penalty, but there is a limit to how much the fine can be increased in economies with limited liabilities. To raise the expected penalty further, one must increase the monitoring probability, and this is costly. We let $c > 0$ denote the cost of increasing monitoring enough to increase the expected penalty by one unit. Thus, if (2) holds, it will bind: there is no reason to monitor so much that (2) holds with strict inequality. Further, if (2) does not hold, then $\theta = 0$: if monitoring is not preventing logging, there is no reason to monitor at all. This implies that for each unit of the forest, either the district protects the unit and ensures that (2) binds, or the district does not protect at all, and that unit of the forest will be cut.

District i has a large forest or resource stock X_i, and it is allowed to monitor each unit of this forest with a different intensity. Since the optimal monitoring intensity for each unit ensures that the expected penalty is either p or 0, it follows that a part of the forest will be protected and conserved, perhaps as a national park, while the remaining part will not be sufficiently protected and thus will eventually be cut. If we once again let x_i denote the extraction level in district i, such that $X_i - x_i$ is the size of the forest that is conserved, then i's payoff is $-cp(X_i - x_i) - v_i x_i$, since $\theta = p$ for the part $(X_i - x_i)$ that is conserved.[18] The model thus suggests that conservation policies will be "place-based" (for example, restricted to geographically limited but protected national parks), as seems to be the case in Indonesia, where "national and provincial governments zone areas of forest land to be logged" (Busch et al., 2015, page 1328).

The combined model. More generally, district i may benefit by the part of the resource that is extracted and sold, x_i, at the same time that it finds it expensive to protect the remaining part, $X_i - x_i$. When the arguments above are combined, the utility of district i becomes:

$$u_i = bpx_i - cp(X_i - x_i) - v_i x_i. \tag{3}$$

Below we define districts as "strong" if they benefit a lot from the sale (b is large) while finding enforcement inexpensive (c is small). We will define districts as "weak" if, instad, b is small while c is large. This terminology is consistent with the literature on state capacity, discussed in the Introduction. It will be convenient to assume that the aggregate resource stock is large enough to serve the entire market:

$$\overline{p} - aX < 0, \text{ where } X \equiv \sum_{i \in N} X_i.$$

Interpretations and generalizations. There are several alternative interpre-

[18] To be precise, let S_i be i's forest stock of size X_i, and let θ_s be the expected penalty when logging unit $s \in S_i$. If the forest units are divisible then i's payoff is

$$u_i = -c \int_{S_i} \theta_s ds - v_i \int_{S_i} 1_s ds,$$

where $1_s = 1$ if $\theta_s < p$ but $1_s = 0$ if $\theta_s \geq p$. Since there will be a corner solution for the optimal θ_s, u_i can be written as $-cp(X_i - x_i) - v_i x_i$.

tations of the combined model such as it is summarized in (3). First, even if all extraction is illegal, the district may have some concern for the welfare of the loggers, in particular if they are poor and/or citizens of the districts. Parameter b may then represent this concern. Alternatively, parameter b may reflect the probability that the government in district i captures the profit from the illegal loggers, even in the areas where the forest is not protected.[19]

The model is simple and can easily be generalized in a number of ways. For example, we allow for district-specific parameters b and c in Section 6. (However, we have concluded that the additional insight does not justify the added complexity.) Without changing the analysis, we can also allow for cross-externalities such that district i loses v_{-i} when the other districts extract, in addition to v_i from i's own extraction. To see that our model already captures this case, suppose that i's true payoff is:

$$\widetilde{u}_i = bpx_i - cp\left(\widetilde{X}_i - x_i\right) - \widetilde{v}_i x_i - \widetilde{v}_{-i} \sum_{j \in N \setminus i} x_j.$$

We can then write the payoff as (3) if we simply define $u_i \equiv \widetilde{u}_i + \overline{p}\widetilde{v}_{-i}/a$ while

$$v_i \equiv \widetilde{v}_i - \widetilde{v}_{-i} \text{ and } X_i \equiv \widetilde{X}_i - \widetilde{v}_{-i}/ca.$$

Thus, a larger cross-externality can be captured by considering a reduction in v_i and X_i in the model described above.

Finally, note that we have linked the districts by assuming that the harvest is sold at a common downstream market, but we could equally well assume that districts hire labor or need inputs from a common upstream market. To see this, suppose that the price of the harvest is fixed at $\widehat{p}$, and consider the wage cost of the labor needed to extract. If the labor supply curve is linear in total supply, and loggers are mobile across districts, then we may write the wage as $\widehat{w} + ax$, where $\widehat{w}$ is a constant and $a > 0$ is the slope of the labor supply curve. Defining $\overline{p} \equiv \widehat{p} + \widehat{w}$, we can write this model as (1)-(3). It is thus equivalent to the model described above.

[19] As a third interpretation, if district i decides to extract x_i^s units for sale in order to raise revenues, such extraction may require infrastructure and roads, which in turn may also lead to illegal logging in the amount αx_i^s, where $\alpha > 0$ measures the amount of illegal logging when the government cuts. Such a complementarity is documented by de Sá et al. (2015). Total extraction is then $x_i = (1+\alpha)\, x_i^s$ even though the fraction of the total profit, captured by the government in district i, is only $b \equiv 1/(1+\alpha)$. The larger the fraction of illegal logging, the smaller b is. Whatever is not cut must be protected, just as before.

3 Equilibrium Extraction

Based on the combined model above, in which resource extraction can be sales-driven or protection-driven, this section discusses the equilibrium amount of extraction and conservation. In particular, we will focus on how equilibrium extraction depends on whether districts are weak or strong and the number of districts; discuss when one district benefits or looses if other districts conserve more; and investigate the effect of political centralization. These results are interesting in themselves, and they are also necessary to describe before we analyze conservation contracts in the next section.

Let $X = \sum_{i \in N} X_i$ be the total size of the resource, while $\overline{v} = \sum_{i \in N} v_i/n$ is the average v_i. It is straightforward to derive the market equilibrium since each u_i is quadratic and concave in x_i, given (1).

Proposition 1. *In equilibrium, extraction is given by:*

$$x_i = \frac{(b+c)\,\overline{p} + acnX_i - ac\sum_{j \in N \setminus i} X_j - nv_i + \sum_{j \in N \setminus i} v_j}{a\,(b+c)\,(n+1)} \Rightarrow \tag{4}$$

$$x = \frac{n}{n+1}\frac{\overline{p}}{a} + \frac{acX - n\overline{v}}{a\,(b+c)\,(n+1)} \Rightarrow \tag{5}$$

$$p = \frac{\overline{p}}{n+1} - \frac{acX - n\overline{v}}{(b+c)\,(n+1)}. \tag{6}$$

We will consider only interior solutions such that the right-hand side of (4) is assumed to be positive but less than X_i for every i. If the right-hand side were instead negative (or larger than X_i), the equilibrium would be $x_i = 0$ (or $x_i = X_i$).

Quite intuitively, aggregate extraction x is larger if demand is large ($\overline{p}/a$ large). But extraction is smaller if the districts' opportunity values are high; and p is then also high. Extraction is also large if c is large while b is small, or if c is small while b is large.[20]

Note that a district i extracts more if its own resource stock is large, since a larger x reduces p and thus the protection cost for the (large) remaining amount, $X_i - x_i$. However, if the other districts are large or have small opportunity costs, then these other districts will extract a lot and this reduces the price p. When p is small, it is

[20] This follows from (5), since $acX > n\overline{v}$ holds only when c is large. Thus, x can be nonmonotonic in state capacity (as measured by a large b and small c): weak states conserve little because they find it expensive, while strong states extract a lot because they profit from it. Acemoglu (2005) finds that countries' performance can be nonmonotonic in state capacity also for other reasons (too high vs. too low taxes).

both less profitable for i to sell the resource, and less expensive for i to protect its resource. For both reasons, a district conserves more when X_j is large or v_j is small, for $j \neq i$.

Having solved for the equilibrium, we can be precise about when districts benefit from a high price. When we take the partial derivative of (3) with respect to p, and substitute with (4), we get:

$$\frac{\partial u_i}{\partial p} = \frac{e_i}{a(n+1)} \text{ where } e_i \equiv b\overline{p} - c(aX - \overline{p}) - nv_i + \sum_{j \in N \setminus i} v_j. \tag{7}$$

With this, it is natural to define our labels in the following precise way.

Definition. *Districts are* strong *and extraction is* sales-driven *if districts benefit from a high price (i.e., $e_i > 0$). Districts are* weak *and extraction is* protection-driven *if districts benefit from a low price (i.e., $e_i < 0$).*

With this definition, districts are strong or, equivalently, extraction is sales-driven if $e_i > 0$, which holds not only when the benefit from profit (b) is large, but also when the market size ($\overline{p}/a$) is large compared to the total resource stock, and when protecting the resource has small costs (c) or low value (v_i). Note that we always have $e_i > 0$ in the standard Cournot model (where $c = 0$) when $x_i > 0$. In contrast, we say that districts are weak and extraction is protection-driven when districts benefit from a low price, since costly monitoring must increase accordingly. This requires that $e_i < 0$, which always holds in the model of illegal extraction (when $b = 0$ and $v_i = v_j$).

Since the price is endogenous and increases when the neighbors conserve, e_i can also be referred to as the intra-district (pecuniary) *externality* from conservation.

Proposition 2. *(i) District i benefits when another district conserves if and only if $e_i > 0$:*

$$\frac{\partial u_i}{\partial(-x_j)} = \frac{e_i}{n+1}.$$

(ii) At the equilibrium conservation levels, we have:

$$u_i = \frac{1}{a(b+c)}\left[\left(\frac{e_i}{n+1}\right)^2 - acv_iX_i\right] \Rightarrow \tag{8}$$

$$sign\frac{\partial u_i}{\partial \overline{p}} = sign\frac{\partial u_i}{\partial v_j} = -sign\,\frac{\partial u_i}{\partial X_j} = sign\; e_i. \tag{9}$$

Part (ii) of Proposition 2 shows that the sign of e_i is also important when evaluating other changes. If the market size $\overline{p}$ increases, the price is higher; a high price is beneficial in a sales-driven model where $e_i > 0$, but not when districts are weak and find protection costly. If a district $j \neq i$ values conservation more, or if j's resource stock is smaller, then j is expected to extract less. District i's utility will then increase if and only if $e_i > 0$.

The sign of e_i is also important for district i's *strategy.* If $e_i > 0$, district i prefers a high price, and thus i has an incentive to keep the price high by strategically conserving more. If $e_i < 0$, district i has an incentive to extract strategically more to keep the price and thus the pressure low.

These strategic incentives are particularly important for a large district which influences the price more by a given change in x_i/X_i. It thus follows that while large districts conserve a *larger* fraction of their resource in a sales-driven model (in order to keep p high), they conserves a *smaller* fraction of their resource when extraction is protection-driven (in order to reduce p and thus the pressure on the resource). This can be seen by inserting (7) into (4) to get:

$$\frac{x_i}{X_i} = \frac{ac}{a(b+c)} + \frac{e_i/X_i}{a(b+c)(n+1)}.$$

Corollary 1. *A larger district i conserves a larger fraction of its resource if and only if $e_i > 0$:*

$$\frac{\partial x_i/X_i}{\partial X_i} = \frac{-e_i/X_i^2}{a(b+c)(n+1)}.$$

The effect of the number of districts, n, is equally ambiguous and interesting. In a sales-driven model, it is well known from Cournot games that if the number of sellers increases, then so does the aggregate quantity supplied, while the price declines. We should thus expect $\partial x/\partial n > 0$ in a sales-driven model. With protection-driven extraction, however, districts conserve *less* when they take into account the fact that the pressure on the resource weakens as a consequence. It is for this reason that large districts conserve less. By inserting (7) into (5), we can see that $\partial x/\partial n < 0$ if and only if $\bar{e} < 0$:

$$x = \frac{cX}{b+c} + \frac{n\bar{e}}{a(b+c)(n+1)}, \text{ where} \tag{10}$$
$$\bar{e} \equiv \frac{1}{n}\sum_{i\in N} e_i = (b+c)\bar{p} - acX - \bar{v}.$$

The number of districts is therefore important. If decision-making authority is centralized, the number of relevant governments n declines while the aggregate resource X remains unchanged. Whether it is only a couple of districts that centralize authority to a common central authority, or all the n districts that centralize power to a single government, we get the following corollary straightforwardly from equation (10).

Corollary 2. *Fix X and $\bar{v}$. Centralization (implying a smaller n) leads to more conservation if districts are strong ($\bar{e} > 0$) but less if districts are weak ($\bar{e} < 0$).*

If authority is centralized to a single central government, C, then $n = 1$ and (10) becomes:

$$x_C = \frac{cX}{b+c} + \frac{\bar{e}}{2a(b+c)} = \frac{\bar{p}(b+c) + acX - \bar{v}}{2a(b+c)}. \tag{11}$$

4 Conservation Contracts

The previous section derived equilibrium conservation as a function of the parameters in the model. In this section, we further assume that every district has a utility function that is linear and additive in money. That is, if $\tau_i \in \Re$ refers to a transfer to district i, then τ_i enters additively in i's objective function.

In particular, we have already suggested that district i's opportunity cost of extraction, v_i, may in part come from lost subsidies or a higher tax on extraction:

$$v_i = v + t_i, \tag{12}$$

where $t_i \in \Re$ can represent an extraction tax, so that the transfer to i would be $\tau_i = -t_i x_i$. We let v be common for the districts, and we define $e \equiv (b+c)\bar{p} - acX - v$.[21]

[21] If, instead, the exogenous parts of v differed for some pair of districts, then a central authority maximizing the sum of utilities would prefer a corner solution where everything is conserved in one district or nothing in the other. Such a corner solution would hinge on the linearity assumptions in our model and is thus uninteresting to emphasize. See, however, Section 6 where we do allow for heterogeneity in the v's, the b's, and the c's.

Given (12), (4) shows that x_i is a function of $\mathbf{t} = (t_1, .., t_n)$:

$$x_i(\mathbf{t}) = \frac{(b+c)\,\overline{p} + ac\,(n+1)\,X_i - acX - v - t_i n + \sum_{j \neq i} t_j}{a\,(b+c)\,(n+1)}. \tag{13}$$

Equations (1)-(3) show that u_i is a function of the vector $\mathbf{x}(\mathbf{t}) = (x_1(\mathbf{t}), ..., x_n(\mathbf{t}))$:

$$u_i^0(\mathbf{x}(\mathbf{t})) = bpx_i(\mathbf{t}) - cp\,(X_i - x_i(\mathbf{t})) - vx_i(\mathbf{t}),$$

where superscript 0 just indicates that the cost of t_i is not taken into account in the definition of u_i^0. Thus, with $\tau_i = -t_i x_i$, i's actual payoff is just as in (3):

$$u_i^0(\mathbf{x}(\mathbf{t})) + \tau_i = bpx_i(\mathbf{t}) - cp\,(X_i - x_i(\mathbf{t})) - (v + t_i)\,x_i(\mathbf{t}).$$

In this section we study contracts between the districts and a principal or a "donor" D. We simply assume that D benefits from conservation and that $u_D = -dx$, where $d > 0$ measures the damage D faces from the districts' extraction. Also D has a quasi-linear function for the payoff $u_D + \tau_D$, where τ_D is the transfer to D. By budget balance, $\tau_D = -\sum_{i \in N} \tau_i$. In the following, we will derive (1) the first-best (Pareto-optimal) allocations, (2) linear contracts between D and a central government C, (3) linear contracts between D and $m \leq n$ districts, and (4) nonlinear contracts.

4.1 The First Best

Since we have assumed transferable utilities and $n+1$ players, any Pareto optimal allocation $\mathbf{x} = (x_1, ..., x_n)$ must maximize $u_D(\mathbf{x}) + \sum_{i \in N} u_i^0(\mathbf{x})$. Pareto optimality cannot pin down the transfers or even the allocation of x_i's when x is given and v is the same for every district, but the Pareto-optimal x is unique.

Proposition 3. *(i) The first-best extraction level is given by:*

$$x_{FB} = \frac{\overline{p}\,(b+c) + acX - v - d}{2a\,(b+c)} = \frac{cX}{b+c} - \frac{d-e}{2a\,(b+c)}. \tag{14}$$

(ii) The first-best x_{FB} is implemented by the decentralized equilibrium if and only if:

$$\bar{t} = t_{FB} \equiv \left(\frac{n+1}{2n}\right) d + \left(\frac{n-1}{2n}\right) e, \text{ where} \tag{15}$$

$$\bar{t} \equiv \sum_{i \in N} t_i / n. \tag{16}$$

Part (i) shows that the expression for x_{FB} equals the expression for x_C if simply $\overline{v}$ in (11) is replaced by $v+d$. Part (ii) of the proposition follows simply from combining (5), (12), and (14). It states that the first-best tax or (subsidy) rate t_{FB} is a weighted average of the two externalities e and d. To understand this, note that even when $d = 0$, $t_i > 0$ is optimal if and only if other districts benefit when i conserves more. This would be the case when districts are strong and extraction sales-driven. When districts are weak and extraction protection-driven, then $t_i < 0$ would be optimal instead.[22]

When decision-making power is centralized to a single authority, then $n = 1$ and the Pigou tax is standard.

Corollary 3. *Under centralization, the first-best is implemented simply by $t_C = d$. Facing $t_C = d$, C will induce its districts to select x_{FB} by, for example $\bar{t} = t_{FB}$.*

The second part of the corollary is just pointing out that since C maximizes the sum of the districts' payoffs, it will tax extraction according to (15) if just d is replaced by t_C. We thus have a formula for how C will implement its desired policy for any given t_C:

$$\bar{t}\,(t_C) \equiv \left(\frac{n+1}{2n}\right) t_C + \left(\frac{n-1}{2n}\right) e.$$

4.2 Contracts under Centralization

While Proposition 3 describes the first best, we now derive the equilibrium contract if D can make a take-it-or-leave-it offer. We assume the extraction level is contractible so that the transfer from D can be a function of $\mathbf{x}$. We start by considering a linear

[22] As a remark on the details, note that the levels of $\partial u_i / \partial(-x_j)$ and e_i depend on t_i. Given (7) and (12), e_i increases in t_j but decreases in t_i and in a common tax t. The Pigou tax that internalizes all externalities is thus

$$t_i = \sum_{j \in N \setminus i} \frac{\partial u_i}{\partial(-x_j)} + d = (n-1)\,\frac{e - n(v+t_i) + \sum_{j \in N \setminus i} t_j}{n+1} + d,$$

which can be written as (15).

contract which takes the form of actual real-world REDD-contracts. If D contracts with C, this means:

$$\tau_C = \max\left\{0, \left(\overline{x}_C - x\right) t_C\right\},$$

where $\overline{x}_C$ is a baseline deforestation level. The contract, which consists of the pair $(t_C, \overline{x}_C)$, implies that C receives t_C dollars for every unit by which the actual extraction x is reduced relative to the baseline level $\overline{x}_C$. If $x \geq \overline{x}_C$, no payment is taking place. When $x < \overline{x}_C$, the transfer can be written as $\tau_C = t_C \overline{x}_C - t_C x$, with the last term being equivalent to a tax t_C, while the first term is equivalent to a lump-sum payment.

If $x < \overline{x}_C$, then C's payoff is $u_C^0\left(x_C\left(t_C\right)\right) + t_C\left(\overline{x}_C - x_C\left(t_C\right)\right)$, where $x_C\left(t_C\right)$ recognizes that x_C is a function of t_C. This function is given by (11), taking into account that $\overline{v} = v + t_C$. Note that x_C is then not a function of the baseline $\overline{x}_C$, which confirms that the $t_C \overline{x}_C$-part of the transfer is like a lump sum.

Since D's objective is to maximize

$$u_D - t_C \cdot \left(\overline{x}_C - x\right), \tag{17}$$

D would prefer to reduce the total transfer τ_C by reducing the baseline $\overline{x}_C$. However, D must ensure that the following incentive constraint for C is satisfied:

$$u_C^0\left(x_C\left(t_C\right)\right) + t_C \cdot \left(\overline{x}_C - x_C\left(t_C\right)\right) \geq u_C^0\left(\widehat{x}_C\right) \forall \widehat{x}_C > \overline{x}_C. \tag{IC$_\text{C}$}$$

That is, C's payoff in equilibrium cannot be smaller than what C could achieve by optimizing as if there were no transfer. In equilibrium, $\overline{x}_C$ will be reduced by D until (IC_C) binds with equality.

Proposition 4. *When D contracts with C, the contract* $(t_C, \overline{x}_C)$ *is:*

$$\begin{aligned} t_C &= d, \\ \overline{x}_C &= x_C(0) - \frac{d}{4a\left(b+c\right)}. \end{aligned} \tag{18}$$

Thus, the optimal rate $t_C = d$ is very simple and independent of the parameters in the model, whether the country is weak or strong, or whether extraction is sales-driven or protection-driven. To derive the result, just substitute (IC_C) into (17), and note that D is induced to maximize the sum $u_C + u_D$.

Corollary 4. *When D contracts with C, the outcome is first best. C faces $t_C = d$ and induces its districts to select x_{FB}, for example by setting taxes satisfying $t = t_{FB}$.*

The baseline $\overline{x}_C$ will be set such that (IC_C) binds and C is exactly indifferent between choosing $x_C(t_C)$ and $x_C(0)$. The indifference means that the benchmark $\overline{x}_C$ will be strictly smaller than the business-as-usual level $x_C(0)$, as illustrated by (18), since otherwise C would strictly benefit from the contract. If $\overline{x}_C$ were not dictated by D, but instead had to equal some historical or business-as-usual level, then D would prefer some other $t_C \neq d$, and the first best would *not* be implemented. This result disproves the typical presumption that the reference level should equal the business-as-usual level.[23]

4.3 Contracts under Decentralization

We now return to the model in which n districts act noncooperatively when deciding on the x_i's. Just as under centralization, we start by considering actual conservation contracts of the form:

$$\tau_i = \max\{0, (\overline{x}_i - x_i)\, t_i\},$$

where $\overline{x}_i$ is the baseline for district i. Suppose D unilaterally designs the contract $(t_i, \overline{x}_i)$ for every $i \in M \subseteq N$, where $m = |M| \leq n$. Even if D would like to contract with all n districts, this may be unfeasible for exogenous (or political) reasons.

Just as under centralization, D must ensure that a district is no worse off in equilibrium where $x_i < \overline{x}_i$ than the district could be by ignoring the contract and picking any other extraction level $\widehat{x}_i > \overline{x}_i$:

$$u_i^0(\mathbf{x}(t)) + t_i \cdot (\overline{x}_i - x_i) \geq u_i^0(\widehat{x}_i, x_{-i}(\mathbf{t}))\; \forall \widehat{x}_i > \overline{x}_i. \qquad (\text{IC}_i)$$

In this incentive constraint, district i takes as given the other districts' extraction levels $x_{-i}(\mathbf{t}) = (x_1(\mathbf{t}), ..., x_{i-1}(\mathbf{t}), x_{i+1}(\mathbf{t}), ...x_m(\mathbf{t}))$ when they expect i to extract $x_i(\mathbf{t})$. When all incentive constraints are satisfied, the extraction levels $\mathbf{x}(t)$ are indeed an equilibrium outcome.

[23] See, for example, Busch et al. (2012) or Dutschke and Angelsen (2008). The latter contribution also discusses why the baseline may be smaller than the business-as-usual (or historical) deforestation level, since this reduces the amount that needs to be paid. This conclusion by Dutschke and Angelsen (2008) is invalid if there are more than one district, we show below (Proposition 5).

However, when the (IC)'s bind, there are multiple equilibria at the extraction stage. If the other districts expect i to extract more, such that $x_i > \overline{x}_i$, then the other districts find it optimal to extract less. This strengthens i's incentive to extract and $x_i > \overline{x}_i$ becomes strictly preferred by i. In this case, i receives no transfers and the outcome is $\mathbf{x}(\mathbf{t}_{-i})$, where $\mathbf{t}_{-i} = (t_1, ..., t_{i-1}, 0, t_{i+1}, ...t_m)$. Of all the multiple equilibria, D prefers the equilibrium in which extraction levels are $\mathbf{x}(t)$. But district $i \in M$ prefers $\mathbf{x}(t)$ to $\mathbf{x}(\mathbf{t}_{-i})$ only if:

$$u_i^0(\mathbf{x}(\mathbf{t})) + t_i \cdot (\overline{x}_i - x_i) \geq u_i^0(\mathbf{x}(\mathbf{t}_{-i})). \quad (\text{PC}_i)$$

If (PC_i) is violated, then i would prefer to reject the contract immediately, by announcing to the other districts that it will not accept any transfers from D. When such a promise is credible, then D must take the participation constraints (PC_i) into account. The problem for D is then to select the m pairs $(t_i, \overline{x}_i)$ in order to maximize:

$$u_D - \sum_{i \in M} t_i \cdot (\overline{x}_i - x_i),$$

subject to the m incentive constraints and the m participation constraints.

Proposition 5. *Suppose D contracts with $m \leq n$ districts.*
(i) When only the (IC_i)'s binds, the optimal contract for D is:

$$\begin{aligned} t_{IC} &= \frac{2}{n+1} d, \\ \overline{x}_i &= x_i(\mathbf{0}) + \frac{4m - 3(n+1)}{4a(b+c)(n+1)} t. \end{aligned} \tag{19}$$

(ii) When only the (PC_i)'s binds, the optimal contract for D is:

$$\begin{aligned} t_{PC} &= \frac{(n+1)d - (n-1)e}{2 + 2m(n-1)}, \\ \overline{x}_i^{PC} &= x_i(\mathbf{0}) + \frac{1}{a(b+c)} \left[\frac{n-1}{(n+1)^2} e + \frac{2n(m-1) - n^2}{(n+1)^2} t \right]. \end{aligned}$$

(iii) Every (IC_i) is strictly stronger than (PC_i) if and only if districts are weak:

$$e < -\left(\frac{4m - n - 3}{4} \right) t. \tag{20}$$

Naturally, when there is only one district ($m = n = 1$), both (i) and (ii) coincide with Proposition 4. To understand part (iii), note that when e and t are large, then district i benefits when the other districts conserve. It is then tempting for i to reject D's offer publicly (rather than simply ignoring it), which implies that (PC) is harder to satisfy than is (IC). Note that a district's size is irrelevant for whether (IC) or (PC) is strongest, as well as for the equilibrium contract, as it is described by Proposition 5.

An attractive outside option to i means that D must increase the baseline $\overline{x}_i$ to ensure that the deal is sufficiently beneficial to i. This is costly to D, but when (PC) binds, D can reduce this cost by reducing t. Consequently, the larger e is, the smaller is the equilibrium t when (PC) binds, as illustrated in part (ii) of the proposition. This argument fails when the binding constraint is instead (IC): At the extraction stage, i can still ignore the offer from D but that will not influence other districts' extraction levels. Such a deviation would thus not be especially beneficial when e is large, so the equilibrium t does not depend on e when (IC) binds.

Note the stark contrast to the first best, which requires t_{FB} to increase in e. Therefore, when e is large, the equilibrium tax is smaller than the first-best tax, and equilibrium extraction is larger than the first-best extraction level. When e is small, the first-best t_{FB} is instead smaller than the equilibrium t. In this case, there is *too much* conservation in equilibrium relative to what is optimal.

Corollary 5. *At the equilibrium contracts, the conservation level is too large compared to the first best if and only if districts are weak (e/d is small):*

$$x < x_{FB} \Leftrightarrow t > t_{FB} \Leftrightarrow \frac{e}{d} < \epsilon, \text{ where}$$

$$\epsilon \equiv \left\{ \begin{array}{lll} -\frac{(n+1)^2-4m}{n^2-1} & \text{if only (IC) binds;} & \text{(i)} \\ -\frac{(n+1)[1+m(n-2)]}{(n-1)(1+mn)} & \text{if only (PC) binds;} & \text{(ii)} \\ -\frac{(4m-n-3)(n+1)}{4m(n+1)-(n-1)(n+3)} & \text{if both (IC) and (PC) bind.} & \text{(iii)} \end{array} \right\}.$$

Example. To illustrate the results with some numbers, consider the situation with two districts, A and B, so $m = n = 2$. In this case, $t_{IC} = 2d/3$ and (PC) does not bind when $e/d < -1/2$. Likewise, $t_{PC} = d/2 - e/6$ and (IC) does not bind for this t when $e/d > -3/7$. When $e/d \in [-1/2, -3/7]$, (IC) and (PC) both bind and t is then given by (20), so $t = -4e/3$. This is all illustrated in Figure 1.

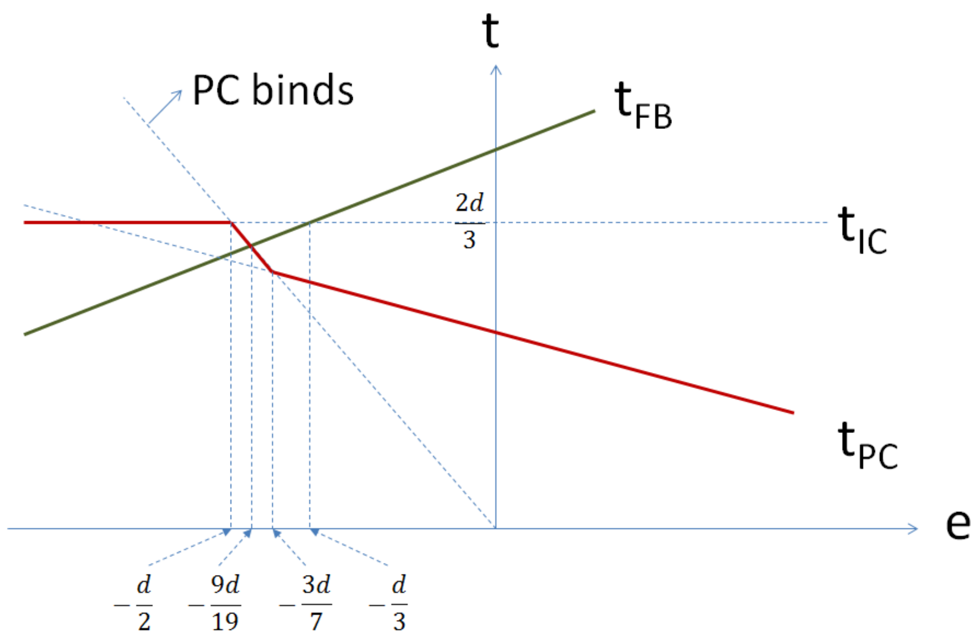

Figure 1: *In our two-district example, there is too little conservation if and only if* $e/d > -9/19$.

In this example, the first best (15) requires that $t_{FB} = 3d/4 + e/4$. In Figure 1, t_{FB} crosses the equilibrium t when $e/d = -9/19 \approx -0.47$, and t_{FB} crosses t_{IC} when $e/d = -1/3$. Thus, when both (IC) and (PC) must be satisfied, there is too little conservation if $e/d > -9/19$. If only (IC) had to be satisfied, there would be too little conservation if $e/d > -1/3$.

4.4 Nonlinear Contracts

So far, we have restricted attention to linear contracts since real-world REDD contracts do take this form. However, such contracts may or may not be optimal from D's point of view. Suppose instead that D could offer transfers that were contingent on the vector of extraction levels, perhaps in a nonlinear way. That is, let D select m functions $\tau_i(\mathbf{x})$. It turns out that while D must still ensure that the incentive constraints are satisfied, every participation constraint can be relaxed.

Proposition 6. *Suppose D can offer any contract* $\tau_i(\mathbf{x}) \geq 0$ *to* $i \in M$.

(i) The incentive constraints are (IC_i) *as before.*

(ii) The participation constraints (PC_i) *can be relaxed so that they are always weaker than the incentive constraints.*

(iii) Consequently, the optimal contract implements x_i*'s and transfers which are iden-*

tical to those of Proposition 5(i).

Thus, when (IC) is in any case the binding constraint, then D cannot do any better than sticking to the linear contracts. This is the case when districts are weak and extraction protection-driven. If, however, districts are strong and extraction sales-driven, then (PC) is the binding constraint and D can do better with nonlinear contracts.[24]

Corollary 6. *From D's perspective, nonlinear contracts are inefficient if and only if (PC) binds at* $t = t_{IC}$*:*

$$\frac{e}{d} > -\frac{4m - n - 3}{2(n+1)}.$$

Condition (PC) can be relaxed in a simple way by offering payments if and only if $x_i = x_i(\mathbf{t}_{IC})$, and zero payment for every other $x_i \neq x_i(\mathbf{t}_{IC})$. Of course, if the districts cannot credibly commit to decline transfers from D, then (PC) would not be binding in the first place. In either case, since (PC)'s can easily be relaxed, we henceforth focus on the case where only (IC)'s bind.

5 Endogenous Regimes

Section 4.2 showed that centralization was first best when the donor could offer conservation contracts, while Section 4.3 showed that decentralized contracts were generically not Pareto optimal. Both sections assumed that the donor contracted with certain districts and governments, and we took their numbers and authority levels to be exogenously given. In some cases, the donor may be able to decide whether it wants to contract with a set of districts independently, or whether it instead wants to contract with their common central government. In other cases, the districts may be capable of centralizing authority, but the incentive to do so may change when the donor is present. A regime change, in turn, may influence conservation. This section (1) studies when the donor would prefer to contract with districts rather than central authorities, (2) endogenizes the political regime and shows when the presence of the donor influences regime change, and (3) describes when the induced regime change increases extraction by more than the contracts themselves reduce it. In the latter case, the presence of the donor reduces conservation.

[24] From Proposition 5 it follows that if t is set according to (19), then (PC) is not binding if (20) holds for that t. This gives Corollary 6.

As a start, consider a subset $L \subseteq M$ containing $l \equiv |L|$ districts. If these districts centralize authority, then l, m, and n all decrease by the same number, denoted by Δ. If L centralizes to a single government, then $\Delta = l - 1$, but we do not require this. We assume that such a regime change does not influence the forest areas over which D can contract. Hence, D contracts with $m - \Delta$ governments after the regime change, while the number of districts without a contract stays unchanged at $n - m$.[25]

We say that L is "large" (relative to N) if:

$$\underline{\epsilon}_L \equiv 1 - \frac{l}{n+1} - \frac{l - \Delta}{n - \Delta + 1} < 0. \quad (21)$$

That is, for L to be large, it is *necessary* that L contains a majority of the decision-making districts *before* centralization $(l > (n+1)/2)$, and it is *sufficient* that L contains a majority *after* decentralization $(l - \Delta > (n - \Delta + 1)/2)$. If L is not large, we say that L is "small."

Our first observation concerns the effect on conservation.

Proposition 7. *If $L \subseteq M$ centralizes, x decreases if and only if e/d is large or M is large, i.e., if:*

$$\frac{e}{d} \geq 2\underline{\epsilon}_M = 2\left(1 - \frac{m}{n+1} - \frac{m - \Delta}{n - \Delta + 1}\right).$$

If M is large (say, $m = n$), then we know from the earlier intuition that centralization reduces x when e/d is large. If M is small, however, a large number $(n - m)$ of other districts will increase x when M reduces x, and thus the condition becomes harder to satisfy. The proposition generalizes Corollary 2 (for the case in which there is no contract, $d = 0$). The proposition also generalizes Corollary 5(i), and the two corollaries coincide when $l = n$ and $\Delta = l - 1$.

5.1 Selecting Contractors

This subsection studies when the donor would prefer contracting with a central government rather than with local governments. If a central government C is already active and regulating the local governments, then C can always undo D's offers to the districts; decentralized contracts would then not be an option for D. If the central government is absent or passive, however, then D may evaluate whether it should con-

[25] In Section 6 we also consider regime change among the districts without a contract.

tract with the districts or instead propose a contract to the union of some districts. The latter option may require central authorities to be activated or created.

As we have already noted, when districts are strong and extraction sales-driven, then a district benefits if the others conserve more, and thus also if the others are offered conservation contracts. These positive externalities are internalized by a benevolent central government maximizing the sum of utilities. When positive externalities are appreciated by the contracting partner, D can extract more of the districts' surplus (by reducing the baseline). Thus, when e is large, D benefits when authority is centralized.

If instead the externality e is small, as when districts are weak and extraction is protection-driven, the argument is reversed. A district then experiences negative externalities when others conserve or sign conservation contracts with D. Negative externalities will be taken into account by central authorities, who will thus reject the contract unless it involves larger transfers. In this case, therefore, D prefers decentralized contracts. This holds even when the first best requires centralization.

Proposition 8. *Decentralized contracts are preferred by D if and only if e/d is small or M is small:*

$$\frac{e}{d} \leq \underline{\epsilon}_M = 1 - \frac{m}{n+1} - \frac{m-\Delta}{n-\Delta+1}.$$

The donor is more likely to prefer decentralized contracts when M is small relative to N, since the other districts will, as a consequence, extract less when e is large. This result also implies that there is a unique number m that maximizes D's payoff (i.e., the second-order conditions w.r.t. m and Δ hold).

A comparison to Proposition 7 is interesting. When M is large, $0 > \underline{\epsilon}_M > 2\underline{\epsilon}_M$. Thus, when $e/d \in (2\underline{\epsilon}_M, \underline{\epsilon}_M)$, D finds having decentralized contracts to be less expensive, even if centralization would have increased conservation. When M is instead small, $0 < \underline{\epsilon}_M < 2\underline{\epsilon}_M$. In this case, when $e/d \in (\underline{\epsilon}_M, 2\underline{\epsilon}_M)$, D finds centralized contracts less expensive, even if decentralization will increase conservation. When e/d is outside these intervals, D prefers the regime that maximizes conservation.

Corollary 7. *(i) If M is large, D always prefers decentralized contracts if this reduces x, but the converse is not true.*
(ii) If M is small, D always prefers centralized contracts if this reduces x, but the converse is not true.

5.2 Equilibrium Regime

While the previous subsection analyzed the regime preferred by the donor, we now study the preferences of the districts and derive the equilibrium regime. In particular, we consider the subset $L \subseteq M$ and investigate when the sum these districts' payoffs is larger if they centralize and lower the number of districts to $l-\ \Delta$.

To understand the following result, consider first the case without the donor. If $L = N$, we know that the sum of payoffs is highest under centralization, since externalities make decentralization inefficient. It is thus intuitive that if L is large, L will prefer centralization in the absence of D. If L is small, however, L may pay more attention to what the *other* districts will do when L decentralizes. If $e > 0$ ($e < 0$), L will conserve less (more) if it decentralizes, and, in response, the other districts will conserve more (less). The effect of the others' actions is beneficial to L (regardless of e). Since the number of other districts is large when L is small, a small L prefers to decentralize in the absence of D. This is confirmed in the following result for the special case where $d \to 0 \Rightarrow |e/d| \to \infty$.

Proposition 9. *(i) If L is large, L prefers decentralization if and only if* $e/d \in [\widehat{\epsilon}_L, \overline{\epsilon}_L]$, *where* $\overline{\epsilon}_L > \widehat{\epsilon}_L > \underline{\epsilon}_M > 2\underline{\epsilon}_M < 0$.
(ii) If L is small, L prefers centralization if and only if $e/d \in [\overline{\epsilon}_L, \widehat{\epsilon}_L]$, *where* $\overline{\epsilon}_L < \widehat{\epsilon}_L < \underline{\epsilon}_M$. *The thresholds are given by:*

$$\begin{aligned}
\widehat{\epsilon}_L &\equiv 1 - \frac{2(m-\Delta)}{n-\Delta+1} - \frac{\frac{2m-\Delta}{n+1} - \frac{2m-2\Delta}{n-\Delta+1}}{\sqrt{\frac{l-\Delta}{l}\left(\frac{n+1}{n-\Delta+1}\right)}+1}, \\
\overline{\epsilon}_L &\equiv 1 - \frac{2(m-\Delta)}{n-\Delta+1} + \frac{\frac{2m-\Delta}{n+1} - \frac{2m-2\Delta}{n-\Delta+1}}{\sqrt{\frac{l-\Delta}{l}\left(\frac{n+1}{n-\Delta+1}\right)}-1}.
\end{aligned}$$

Part (i) shows that a large L may prefer to decentralize authority in the presence of an important D (i.e., unless d is very small). As revealed by the inequality $\widehat{\epsilon}_L > \underline{\epsilon}_M$, this will be the case only when e is so large that D would have preferred centralized contracts. Thus, L decentralizes only when this harms D.

Part (ii) similarly states that $\widehat{\epsilon}_L < \underline{\epsilon}_M$. Taken together with Proposition 8, this implies that whenever a small L prefers to centralize authority, then D would instead have preferred decentralized contracts. It may come as no surprise that D and L have conflicting preferences, given that the regime influences the transfers from D.

Corollary 8. *(i) Suppose L is large. If L prefers decentralization, D prefers centralized contracts. If D prefers decentralized contracts, L prefers centralization.*
(ii) Suppose L is small. If L prefers centralization, D prefers decentralized contracts. If D prefers centralized contracts, L prefers decentralization.

Another corollary of Proposition 9(i) can be drawn from the statement $\widehat{\epsilon}_L > 2\underline{\epsilon}_M$. Together with Proposition 7, this implies that whenever a large L prefers decentralization, then decentralization reduces conservation. A related fact can be derived from part (ii) for the special case in which $L = M$. In that case, we know that for a small L, $\underline{\epsilon}_M > 0$ and thus $\widehat{\epsilon}_L < \underline{\epsilon}_M < 2\underline{\epsilon}_M$. Hence, whenever a small $L = M$ prefers centralization, centralization reduces conservation.

Corollary 9. (i) *Suppose L is large. If L prefers decentralization, decentralization increases extraction. If decentralization reduces extraction, then L prefers centralization.*
(ii) Suppose L=M is small. If L prefers centralization, centralization increases extraction. If centralization reduces extraction, then L prefers decentralization.

5.3 The Donor's Influence on Conservation

In the absence of the donor (or when $d \to 0$), Proposition 9 states that a large L would centralize while a small L would decentralize. Corollary 8 states further that if a large L decentralizes, D would have preferred that it didn't; and if a small L centralizes, D would have preferred that it didn't. We can summarize these observations as follows.

Corollary 10. *The presence of the donor may induce a regime change. If so, the induced regime change always harms the donor.*

Corollary 9 states further that if L is large, or if if L and M are small, then the induced regime change leads to less conservation. Based on these findings, one may question whether the reduced conservation levels following regime change can outweigh the effect of the contracts offered by D. If so, the very presence of D leads to regime change and so much more extraction that a larger part of the resource would have been conserved if D, as well as D's contracts, had been absent. In this case, D's presence does more harm than good and D would have preferred to commit to abstaining from offering contracts, if such a commitment were feasible.

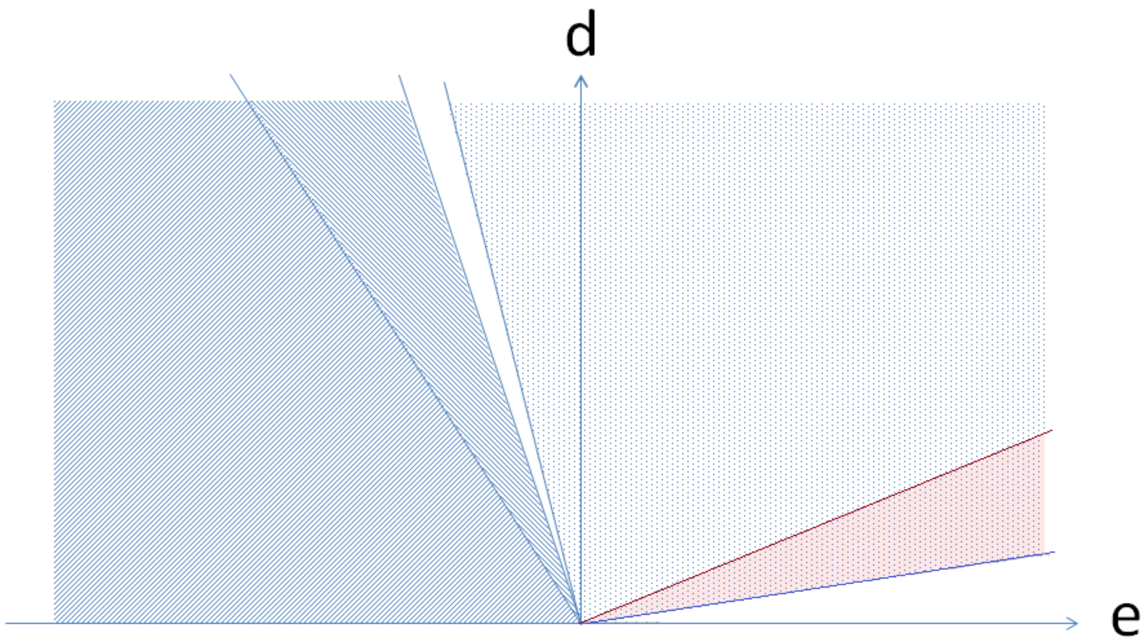

Figure 2: *Even if centralization leads to the first best, the donor prefers decentralized contracts when e and d are small (shaded area), while the districts prefer decentralization when e and d are large (dotted area). In the colored dotted area, the regime change raises extraction by more than the contracts reduce it. Lines are drawn for our two-district example.*

Proposition 10. *(i) Suppose L is large. If $e/d \in [\widehat{\epsilon}_L, \overline{\epsilon}_L]$, the presence of D induces decentralization and, despite the contracts, x increases when the following holds:*

$$\frac{e}{d} > \widetilde{\epsilon}_L \equiv \frac{2m\,(n-\Delta+1)}{(n+1)\,\Delta}.$$

(ii) Suppose L is small. If $e/d \in [\overline{\epsilon}_L, \widehat{\epsilon}_L]$, the presence of D induces centralization and, when also $e/d < -\widetilde{\epsilon}_L$, x increases, despite the contracts.

Consider again our example with two districts. Decentralized contracts are preferred by D in the shaded area in Figure 2, where $e/d < \underline{\epsilon}_M = -1/6 \approx -0.17$, even though decentralization reduces conservation when $e/d > 2\underline{\epsilon}_M = -1/3$ (where the shaded area has downward-sloping lines). The districts, however, prefer decentralization only when they are stronger and $e/d \in (\widehat{\epsilon}_L, \overline{\epsilon}_L) \approx (-0.16,\ 5.5)$, i.e., in the dotted area. Furthermore, note that $\widetilde{\epsilon}_L = 8/3 \in (\widehat{\epsilon}_L, \overline{\epsilon}_L)$. Thus, for every $e/d \in (8/3,\ 5.5)$, which corresponds to the colored and dotted area, the presence of D motivates the districts to decentralize and the accompanying increase in x outweighs the effect of the contracts. The donor's presence is then reducing conservation.

6 Extensions

The workhorse model above is simple and tractible, and it can thus be generalized in several ways. We here report on two such extensions.

6.1 Heterogeneity

A strong assumption above has been that districts are symmetric in that the parameters b and c have been the same for all districts. Thus, all districts have been strong, or all have been weak. This condition simplified our analysis, without removing the results we have wanted to emphasize. The symmetry assumption are not driving our results, however, and our main results hold also when we allow for more asymmetry.

Proposition 11. *Suppose parameters b, c, and v, may all differ across the districts (in addition to the resource stock sizes X_i).*
(i) District i extracts more if c_i is large while b_i is small, but less if c_j is large or b_j is small, when $j \neq i$:

$$x_i = \frac{\overline{p}}{a(n+1)} + \frac{n}{n+1}\frac{ac_iX_i - v_i - t_i}{a(b_i + c_i)} - \frac{1}{(n+1)}\sum_{j \in N \backslash i}\frac{ac_jX_j - v_j - t_j}{a(b_j + c_j)}$$

(ii) When only incentive constraints bind, the optimal contract is characterized by (19) in Proposition 5.

It is easy to confirm that part (i) of the proposition boils down to (4) when b_i and c_i are the same across districts. Despite these generalizations, we have concluded that the added complexity of allowing for asymmetry is not justified by the added insight.

6.2 Regime Change Among Nonparticipating Districts

While we have above considered only regime changes among the districts which D contracts with, it may also be the case that the districts that are not contracted with decides to centralize or decentralize, in order to influence the contract which D offers to the other districts. The proof for this mechanism can be found in the Appendix:

Proposition 12. *(i) If L is large, L prefers decentralization if and only if $e/d \in [\widehat{\epsilon}_L, \overline{\epsilon}_L]$, where the two tresholds satisfy $\widehat{\epsilon}_L < \overline{\epsilon}_L < 0$.*

(ii) If L is small, L prefers centralization if and only if $e/d \in [\overline{\epsilon}_L, \widehat{\epsilon}_L]$, *where now* $\overline{\epsilon}_L < \widehat{\epsilon}_L$ *and* $\overline{\epsilon}_L < 0$.

Naturally, the condition coincides with (29) when $d \to 0$ or $|e/d| \to \infty$, since the donor is then unimportant. The reason why L may want to (de)centralize in the presence of the donor is that a regime change in L will influence what the districts that are contracted with will do, and the benefit of this will depend on what they will do in the absence of a regime change (and that, in turn, depends on d). A large L may want to decentralize because this will reduce extraction in the other districts (when e is small), and this can be good for L if extraction is sales-driven in L because the contracted-with districts are extracting little (because d is large). Similarly, a small L may want to centralize when e and d are large, since this regime change lead to less extraction in the other districts and this is beneficial when d is large.

7 Conclusions and Policy Lessons

This paper presents a novel and tractable model of conservation. We allow for many districts and recognize that extracting some of the resource increases the harvest supply and thus decreases the price and the monitoring costs for the places that are to be conserved. The externality from one district's conservation on others can be positive or negative, depending on state capacities and the size of the resource stock. The model can be used to study various types of resources and alternative motivations for extractions, but it is motivated in particular by deforestation in the tropics.

The analysis generates several policy lessons, for example regarding how decentralization of authority influences conservation. If districts are "strong" and extraction is sales-driven, then districts extract too much since they do not internalize the effect on other districts' profit. A transfer of authority from the local to the federal level will then lead to more conservation and less extraction. If districts are "weak" and extraction is driven by the high cost of protection, then districts might conserve too much since protection in one district can increase the pressure to extract in neighboring districts. In this case, centralizing authority will reduce conservation and increase extraction. These results may also help to explain the mixed empirical evidence: as discussed in the Introduction, decentralization has increased deforestation in Indonesia, while it has reduced deforestation in other areas, such as the Himalayas.

We employ the model to analyze how the optimal conservation contract depends on local institutions and the driver of extraction. Under centralization to a single government, simple Pigou-like contracts are optimal and first best. With several independent districts, however, linear contracts can lead to too much conservation when districts are weak, while they are outperformed by nonlinear contracts when districts are strong. We also show that the donor benefits from contracting with districts if these are weak and extraction is illegal, but with central authorities if districts are strong and extraction sales-driven. These policy lessons are important when designing real-world conservation (or REDD) contracts.

Policy-makers should also be alert to our finding that institutions may be altered by the donor's presence. The districts may prefer to centralize authority if they are weak, and to decentralize if they are strong. A regime change that is induced by the donor's presence (whether this means decentralization or centralization) always harms the donor. Furthermore, donor-induced regime change is likely to increase extraction, and this increase can be so large that it outweighs the effect of the contracts themselves. In these cases, it is essential that the donor builds a reputation for only contracting with pre-specified authority levels.

Our workhorse model is tractable and can be extended in several directions. Other scholars may want to take advantage of this tractability, since the benchmark results we have derived rely on a number of limiting assumptions. In particular, future research should allow the resource (whether renewable or exhaustible) to be extracted over time in a dynamic setting, the functional forms ought to be generalized, parameters might be privately known, and the outcome may also be stochastic. Allowing for these and other generalizations are necessary to further improve our understanding of how the world's natural resources can be conserved.

References

Acemoglu, Daron, "Politics and Economics in Weak and Strong States," *Journal of Monetary Economics*, 2005, *52* (7), 1199–1226.

_ **, Camilo García-Jimeno, and James A. Robinson**, "State Capacity and Economic Development: A Network Approach," *American Economic Review*, 2015, *Forthcoming.*

Alemagi, Dieudonne and Robert A. Kozak, "Illegal Logging in Cameroon: Causes and the Path Forward," *Forest Policy and Economics*, 2010, *12* (8), 554–561.

Alston, Lee J. and Krister Andersson, "Reducing Greenhouse Gas Emissions by Forest Protection: the Transaction Costs of REDD," *Climate Law*, 2011, *2* (2), 281–289.

Amacher, Gregory S., Erkki Koskela, and Markku Ollikainen, "Royalty Reform and Illegal Reporting of Harvest Volumes Under Alternative Penalty Schemes," *Environmental and Resource Economics*, 2007, *38* (2), 189–211.

_ **, Markku Ollikainen, and Erkki Koskela**, "Corruption and Forest Concessions," *Journal of Environmental Economics and Management*, 2012, *63* (1), 92–104.

Angelsen, Arild, "Playing Games in the Forest: State-local Conflicts of Land Appropriation," *Land Economics*, 2001, *77* (2), 285–299.

_ , *Moving Ahead with REDD: Issues, Options and Implications*, CIFOR, 2008.

_ **and David Kaimowitz**, "Rethinking Causes Lessons from of the Models Deforestation : Lessons from Economic Models," *The World Bank Research Observer*, 1999, *14* (1), 73–98.

Antinori, Camille and Gordon Rausser, "Collective Choice and Community Forestry Management in Mexico: An Empirical Analysis," *Journal of Development Studies*, 2007, *43* (3), 512–536.

Atmadja, Stibniati and Louis Verchot, "A Review of the State of Research, Policies and Strategies in Addressing Leakage from Reducing Emissions from Deforestation and Forest Degradation (REDD+)," *Mitigation and Adaptation Strategies for Global Change*, 2012, *17* (3), 311–336.

Avenhaus, Rudolf, Bernhard Von Stengel, Shmuel Zamir, Robert Aumann, and Sergiu Hart, "Inspection Games," *Handbook of Game Theory with Economic Applications*, 2002, *3*, 1947–1987.

Baland, Jean-Marie, Pranab Bardhan, Sanghamitra Das, and Dilip Mookherjee, "Forests to the People: Decentralization and Forest Degradation in the Indian Himalayas," *World Development*, 2010, *38* (11), 1642–1656.

Besley, Timothy and Torsten Persson, "The Origins of State Capacity: Property Rights, Taxation and Policy," *The American Economic Review*, 2009, *99* (4), 1218–1244.

_ **and** _ , *Pillars of Prosperity: The Political Economics of Development Clusters*, Princeton, 2011.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber, "The Political Economy of Deforestation in the Tropics," *The Quarterly Journal of Economics*, 2012, *127* (4), 1707–1754.

Busch, Jonah, Kalifi Ferretti-Gallon, Jens Engelmann, Max Wright, Kemen G. Austin, Fred Stolle, Svetlana Turubanova, Peter V. Potapov, Belinda Margono, Matthew C. Hansen et al., "Reductions in Emissions from

Deforestation from Indonesia's Moratorium on New Oil Palm, Timber, and Logging Concessions," *Proceedings of the National Academy of Sciences*, 2015, *112* (5), 1328–1333.

_ , **Ruben N. Lubowski, Fabiano Godoy, Marc Steininger, Arief A. Yusuf, Kemen Austin, Jenny Hewson, Daniel Juhn, Muhammad Farid, and Frederick Boltz**, "Structuring Economic Incentives to Reduce Emissions from Deforestation within Indonesia," *Proceedings of the National Academy of Sciences*, 2012, *109* (4), 1062–1067.

Chhatre, Ashwini and Arun Agrawal, "Forest Commons and Local Enforcement," *Proceedings of the National Academy of Sciences*, 2008, *105* (36), 13286–13291.

Chiroleu-Assouline, Mireille, Jean-Christophe Poudou, and Sébastien Roussel, "North/South Contractual Design through the REDD+ Scheme," *FEEM Working Paper*, 2012, (89), 106–112.

Clarke, Harry R, William J Reed, and Ram M Shrestha, "Optimal Enforcement of Property Rights on Developing Country Forests Subject to Illegal Logging," *Resource and Energy Economics*, 1993, *15* (3), 271–293.

de Sá, Saraly Andrade, Philippe Delacote, and Eric Nazindigouba Kéré, "Spatial Interactions in Tropical Deforestation: An Application to the Brazilian Amazon," Mimeo, 2015.

Delacote, Philippe, "How Concessions' Size may Influence Systemic Corruption in Forest Harvesting: A Theoretical Assessment," *Working Paper - Cahiers du LEF 2011-05, Laboratoire d'Economie Forestiere, AgroParisTech-INRA.*, 2011.

_ **and Gabriela Simonet**, "Readiness and Avoided Deforestation Policies: on the use of the REDD Fund," *Working Paper 1312, Chaire Economie du Climat.*, 2013.

Dokken, Therese, Susan Caplow, Arild Angelsen, and William D Sunderlin, "Tenure issues in REDD+ Pilot Project Sites in Tanzania," *Forests*, 2014, *5* (2), 234–255.

Dutschke, Michael and Arild Angelsen, "How do we Ensure Permanence and Assign Liability?," in Arild Angelsen, ed., *Moving ahead with REDD: Issues, Options and Implications*, Center for International Forestry Research (CIFOR), 2008, chapter 8, pp. 77–86.

Economist, "Money can Grow on Trees," *The Economist*, September 2010.

Edenhofer, O, R Pichs-Madruga, Y Sokona, E Farahani, S Kadner, K Seyboth, A Adler, I Baum, S Brunner, P Eickemeier et al., "Climate Change 2014: Mitigation of Climate Change: Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change," *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*, 2014, *1.*

Eeckhout, Jan, Nicola Persico, and Petra E Todd, "A Theory of Optimal Random Crackdowns," *The American Economic Review*, 2010, *100* (3), 1104–1135.

Engel, Stefanie, Stefano Pagiola, and Sven Wunder, "Designing Payments for Environmental Services in Theory and Practice: An Overview of the Issues," *Ecological Economics*, 2008, *65* (4), 663–674.

Epple, Dennis and Thomas Nechyba, "Fiscal Decentralization," *Handbook of Regional and Urban Economics*, 2004, *4*, 2423–2480.

Foster, Andrew D. and Mark R. Rosenzweig, "Economic growth and the Rise of Forests," *The Quarterly Journal of Economics*, 2003, *118* (2), 601–637.

Gaudet, Gérard and Stephen W. Salant, "Increasing the Profits of a Subset of Firms in Oligopoly Models with Strategic Substitutes," *The American Economic Review*, 1991, *81* (3), 658–665.

Genicot, Garance and Debraj Ray, "Contracts and Externalities: How Things Fall Apart," *Journal of Economic Theory*, 2006, *131* (1), 71–100.

Gjertsen, Heidi, Theodore Groves, David A. Miller, Eduard Niesten, Dale Squires, and Joel Watson, "A Contract-theoretic Model of Conservation Agreements," *Department of Economics, University of California San Diego*, 2010.

Gonzalez-Navarro, Marco, "Deterrence and geographical externalities in auto theft," *American Economic Journal: Applied Economics*, 2013, *5* (4), 92–110.

Hansen, Matthew C, Peter V Potapov, Rebecca Moore, Matt Hancher, SA Turubanova, Alexandra Tyukavina, David Thau, SV Stehman, SJ Goetz, TR Loveland et al., "High-resolution Global Maps of 21st-century Forest Cover Change," *Science*, 2013, *342* (6160), 850–853.

Harris, Nancy L, Sandra Brown, Stephen C Hagen, Sassan S Saatchi, Silvia Petrova, William Salas, Matthew C Hansen, Peter V Potapov, and Alexander Lotsch, "Baseline Map of Carbon Emissions from Deforestation in Tropical Regions," *Science*, 2012, *336* (6088), 1573–1576.

Harstad, Bård, "Buy Coal! A Case for Supply-side Environmental Policy," *Journal of Political Economy*, 2012, *120* (1), 77–115.

_ , "The Market for Conservation and Other Hostages," Mimeo, 2015.

Hartwick, John M., Ngo Van Long, and Huilan Tian, "Deforestation and Development in a Small Open Economy," *Journal of Environmental Economics and Management*, 2001, *41* (3), 235–251.

Helsley, Robert W., "Urban Political Economics," *Handbook of Regional and Urban Economics*, 2004, *4*, 2381–2421.

Jacobson, Sarah, "Temporal spillovers in Land Conservation," *Journal of Economic Behavior & Organization*, 2014, *107*, 366–379.

Johnson, Shane D., Rob T. Guerette, and Kate J. Bowers, "Crime Displacement and Diffusion of Benefits," in David P. Farrington Brandon C. Welsh, ed., *The Oxford Handbook of Crime Prevention*, Oxford University Press, 2012, chapter 17, pp. 337–353.

Kaimovitz, David, Cristian Vallejos, Pablo B. Pacheco, and Raul Lopez, "Municipal Governments and Forest Management in Lowland Bolivia," *The Journal of Environment & Development*, 1998, *7* (1), 45–59.

Kaimowitz, David and Arild Angelsen, *Economic Models of Tropical Deforestation: A Review*, Cifor, 1998.

Kamien, Morton I. and Israel Zang, "The Limits of Monopolization through Acquisition," *The Quarterly Journal of Economics*, 1990, *105* (2), 465–499.

Karsenty, Alain, "The Architecture of Proposed REDD Schemes after Bali: Facing Critical Choices," *International Forestry Review*, 2008, *10* (3), 443–457.

Kerr, Suzi C., "The economics of International Policy Agreements to Reduce Emissions from Deforestation and Degradation," *Review of Environmental Economics and Policy*, 2013, *7* (1), 47–66.

Kremer, Michael and Charles Morcom, "Elephants," *American Economic Review*, 2000, *90* (1), 212–234.

Lando, Henrik and Steven Shavell, "The Advantage of Focusing Law Enforcement Effort," *International Review of Law and Economics*, 2004, *24* (2), 209–218.

Mason, Charles F., "Optimal Contracts for Discouraging Deforestation with Risk Averse Agents," Mimeo, 2013.

_ **and Andrew J. Plantinga**, "The Additionality Problem with Offsets: Optimal Contracts for Carbon Sequestration in Forests," *Journal of Environmental Economics and Management*, 2013, *66* (1), 1–14.

McAllister, Ryan, Rodney Beard, and John Asufa-Adjaye, "An Optimal Control Model of Deforestation in the PDR of Laos," *44th Conference of Australian Agricultural and Resource Economics Society, Sydney, Australia*, 2000, (123704).

Mehlum, Halvor, Karl Moene, and Ragnar Torvik, "Institutions and the Resource Curse," *The Economic Journal*, 2006, *116* (508), 1–20.

Mendelsohn, Robert, "Property Rights and Tropical Deforestation," *Oxford Economic Papers*, 1994, *46*, 750–756.

Meyfroidt, Patrick and Eric F. Lambin, "Forest Transition in Vietnam and Displacement of Deforestation Abroad," *Proceedings of the National Academy of Sciences*, 2009, *106* (38), 16139–16144.

Murray, Brian C., "Leakage from an Avoided Deforestation Compensation Policy: Concepts, Empirical Evidence, and Corrective Policy Options," *Nicholas Institute for Environmental Policy Solutions, Duke University, Durham, NC.*, 2008.

_ **, Bruce A. McCarl, and Heng-Chi Lee**, "Estimating Leakage from Forest Carbon Sequestration Programs," *Land Economics*, 2004, *80* (1), 109–124.

Oates, Wallace E., *Fiscal Federalism*, Harcourt Brace Jovanovich., 1972.

Parker, C., M. Mitchell A .and Trivedi, and N. Mardas, *The Little REDD+ Book. Global Canopy Foundation*, Oxford, UK, 2009.

Perry, Martin K. and Robert H. Porter, "Oligopoly and the Incentive for Horizontal Merger," *The American Economic Review*, 1985, *75* (1), 219–227.

Phelps, Jacob, Edward E. L. Webb, and Arun Agrawal, "Does REDD+ Threaten to Recentralize Forest Governance," *Science*, 2010, *328* (5976), 312–313.

Polinsky, A. Mitchell and Steven Shavell, "The Theory of Public Enforcement of Law," *Handbook of Law and Economics*, 2007, *1*, 403–454.

Ribot, Jesse C., *African Decentralization: Local Actors, Powers and Accountability*, UNRISD Geneva, 2002.

Robinson, Elizabeth J. and Razack B. Lokina, "A Spatial–temporal Analysis of the Impact of Access Restrictions on Forest Landscapes and Household Welfare in Tanzania," *Forest Policy and Economics*, 2011, *13* (1), 79–85.

_ **and** _ , "Efficiency, Enforcement and Revenue Tradeoffs in Participatory Forest Management: An Example from Tanzania," *Environment and Development Economics*, 2012, *17* (01), 1–20.

_ **, Heidi J. Albers, Charles Meshack, and Razack B. Lokina**, "Implementing REDD through Community-based Forest Management: Lessons from Tanzania," *Natural Resources Forum*, 2013, *37* (3), 141–152.

Robinson, James A., Ragnar Torvik, and Thierry Verdier, "Political Foundations of the Resource Curse," *Journal of Development Economics*, 2006, *79* (2), 447–468.

Robinson, James A, Ragnar Torvik, and Thierry Verdier, "Political Foundations of the Resource Curse: A Simplification and a Comment," *Journal of Development Economics*, 2014, *106*, 194–198.

Salant, Stephen W., Sheldon Switzer, and Robert J. Reynolds, "Losses from Horizontal Merger: the Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium," *The Quarterly Journal of Economics*, 1983, *98* (2), 185–199.

Segal, Ilya, "Contracting with Externalities," *The Quarterly Journal of Economics*, 1999, *114* (2), 337–388.

Shyamsundar, Priya and Rucha Ghate, "Rights, Rewards, and Resources: Lessons from Community Forestry in South Asia," *Review of Environmental Economics and Policy*, 2014, *8* (1), 80–102.

Somanathan, E., R. Prabhakar, and Bhupendra Singh Mehta, "Decentralization for Cost-effective Conservation," *Proceedings of the National Academy of Sciences*, 2009, *106* (11), 4143–4147.

Stocker, T. F., D. Qin, GK Plattner, M Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, B. Bex, and B. M. Midgley, *IPCC, 2013: Climate Change 2013: the Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, 2013.

Torvik, Ragnar, "Why Do Some Resource-abundant Countries Succeed while Others Do Not?," *Oxford Review of Economic Policy*, 2009, *25* (2), 241–256.

van der Ploeg, Frederick, "Natural Resources: Curse or Blessing?," *Journal of Economic Literature*, 2011, *49* (2), 366–420.

8 Appendix: Proofs

Proof of Proposition 1.
Note that the first-order condition when maximizing (3) w.r.t. x_i and subject to (1) gives:

$$
\begin{aligned}
x_i &= \frac{p}{a} + \frac{caX_i - v_i}{a(b+c)} && (22)\\
&= \frac{\overline{p} - ax_{-i}}{2a} + \frac{caX_i - v_i}{2a(b+c)}, && (23)
\end{aligned}
$$

if the right-hand side is in $[0, X_i]$. The second-order condition trivially holds. By summing over the x_i's as given by (22) and combining that sum with (1), we get (5) and (6), and by inserting (6) into (22), we get (4). □

Proof of Proposition 2.
(i) From (3) we immediately get (when $j \neq i$ and using the Envelope theorem):

$$
\begin{aligned}
\frac{\partial u_i}{\partial x_j} &= \frac{\partial u_i}{\partial p}\frac{\partial p}{\partial x_j} = -a\left[(b+c)x_i - cX_i\right]\\
&= -\frac{(b+c)\overline{p} - acX - v_i n + \sum_{j\neq i} v_j}{n+1},
\end{aligned}
$$

when we substitute in for (4). With (7), we can write $\partial u_i/\partial x_j = -e_i/(n+1)$.
(ii) When we combine (7) with (4) and (6), we get:

$$
\begin{aligned}
x_i &= \frac{e_i}{a(b+c)(n+1)} + \frac{cX_i}{b+c} \text{ and}\\
p &= \frac{e_i}{(b+c)(n+1)} + \frac{v_i}{b+c}.
\end{aligned}
$$

Thus, we can write (3) as:

$$
\begin{aligned}
u_i &\equiv x_i\left((b+c)p - v_i\right) - pcX_i\\
&= \left(\frac{e_i}{a(b+c)(n+1)} + \frac{cX_i}{b+c}\right)\frac{e_i}{n+1} - \left(\frac{e_i}{(b+c)(n+1)} + \frac{v_i}{b+c}\right)cX_i,
\end{aligned}
$$

which can be written as (8). Given (7), differentiating (8) gives (9). □

Proof of Proposition 3.
(i) Since $u_D(\mathbf{x}) + \sum_{i\in N} u_i^0(\mathbf{x}) = bpx - cp(X-x) - vx - dx$, the f.o.c. when maximizing w.r.t. x can be written as (14). The second-order condition trivially holds.
(ii) With (12), we can write (10) as

$$
x = \frac{ac}{a(b+c)}X + \frac{ne - \sum_i t_i}{a(b+c)(n+1)}. \tag{24}
$$

This x equals x_{FB} if and only if (15) holds. □

Proof of Proposition 4.
For a given t_C, D prefers to reduce $\overline{x}_C$ as much as possible, so (IC_C) will bind. Solving (IC_C) for $(\overline{x}_C - x)t_C$ and inserting that term into (17), we note that D's objective is to

maximize $-dx_C(t_C)+u_C^0(x_C(t_C))-u_C^0(\widehat{x}_C)=u_D(x_C(t_C))+u_C^0(x_C(t_C))-u_C^0(\widehat{x}_C)$. D is thus maximizing the sum of payoffs (since $-u_C^0(\widehat{x}_C)$ is independent of t_C), implying the same outcome as in the first best: $x_C=x_{FB}$ and $t_C=d$.

To derive $\overline{x}_C$, note that we can rewrite a binding (IC_C) to:

$$t_C\overline{x}_C=u_C^0(\widehat{x}_C)-\left[u_C^0(x_C(t_C))-t_Cx_C\right], \tag{25}$$

where both $u_C^0(\widehat{x}_C)$ and the bracket follow from (8), and with (7) and (12), e_i is replaced by e when C ignores the contract while otherwise $e_i=e-t_C$. Thus, we can write (25) as:

$$t_C\overline{x}_C=\frac{1}{a(b+c)}\left[\frac{e^2}{4}-cavX-\frac{(e-t_C)^2}{4}+ca(v+t_C)X\right],$$

which can be rewritten as (18) when $t_C=d$. □

Proof of Proposition 5.

The proof starts by deriving $\max_{\widehat{x}_i}u_i^0(\widehat{x}_i,x_{-i}(\mathbf{t}))$. From (23) and (4), we find i's optimal response to $x_{-i}(\mathbf{t})$, if i decided to ignore the contract:

$$x_i^I=\frac{\overline{p}-ax_{-i}(\mathbf{t})}{2a}+\frac{caX_i-v}{2a(b+c)}=x_i+\frac{t_i}{2a(b+c)},$$

where x_i is given by (13). This results in a price

$$p^I=p-\frac{t_i}{2(b+c)},$$

where $p=\overline{p}-a\sum_i x_i(\mathbf{t})$. Thus,

$$\begin{aligned}
u_i^0(\widehat{x}_i,x_{-i}(\mathbf{t})) &= \left[(b+c)p^I-v\right]x_i^I-p^IcX_i\\
&= \left[(b+c)\left(p-\frac{t_i}{2(b+c)}\right)-v\right]\left(x_i+\frac{t_i}{2a(b+c)}\right)-\left(p-\frac{t_i}{2(b+c)}\right)cX_i\\
&= u_i^0(\mathbf{x}(\mathbf{t}))+\left[(b+c)\left(p-\frac{t_i}{2(b+c)}\right)-v\right]\frac{t_i}{2a(b+c)}-\frac{t_i}{2}x_i+\frac{cX_it_i}{2(b+c)}\\
&= u_i^0(\mathbf{x}(\mathbf{t}))+\left[(b+c)p-\frac{t_i}{2}-v-a(b+c)x_i+acX_i\right]\frac{t_i}{2a(b+c)}\\
&= u_i^0(\mathbf{x}(\mathbf{t}))+\frac{t_i^2}{4a(b+c)},
\end{aligned} \tag{26}$$

when we use (22). With this, (IC_i) boils down to

$$\tau_i\geq u_i^0(\widehat{x}_i,x_{-i}(\mathbf{t}))-u_i^0(\mathbf{x}(\mathbf{t}))=\frac{t_i^2}{4a(b+c)}. \tag{27}$$

(i) When only the IC's bind, D maximizes

$$
\begin{aligned}
& u_D + \sum_{i\in M} \left[u_i^0 \left(\mathbf{x}\left(t\right)\right) - u_i^0 \left(\widehat{x}_i, x_{-i}\left(\mathbf{t}\right)\right)\right] \\
= & -d\left[\frac{ac}{a\left(b+c\right)}X + \frac{ne - \sum_i t_i}{a\left(b+c\right)\left(n+1\right)}\right] - \sum_{i\in M} \frac{t_i^2}{4a\left(b+c\right)}.
\end{aligned}
$$

For each t_i, $i \in M$, the first-order condition becomes

$$
\frac{d}{a\left(b+c\right)\left(n+1\right)} - \frac{t_i}{2a\left(b+c\right)} = 0,
$$

giving (i). The second-order condition trivially holds.

To find $\overline{x}_i^{IC}$, rewrite a binding (27) to:

$$
\begin{aligned}
\tau_i &= t_i\left(\overline{x}_i - x_i\right) = \frac{t_i^2}{4a\left(b+c\right)} \Leftrightarrow \\
\overline{x}_i &= \frac{t_i}{4a\left(b+c\right)} + x_i\left(\mathbf{0}\right) - \frac{t_i n - \sum_{j\neq i} t_j}{a\left(b+c\right)\left(n+1\right)} = x_i\left(\mathbf{0}\right) + \frac{4m - 3\left(n+1\right)}{4a\left(b+c\right)\left(n+1\right)}t \\
&= x_i^I - \frac{t_i}{4a\left(b+c\right)} = \frac{x_i + x_i^I}{2}.
\end{aligned}
$$

(ii) Note that (PC_i) can be rewritten as:

$$
t_i\overline{x}_i \geq u_i^0\left(\mathbf{x}\left(\mathbf{t}_{-i}\right)\right) - \left[u_i^0\left(\mathbf{x}\left(\mathbf{t}\right)\right) - t_i x_i\right]
$$

where both $u_i^0\left(\mathbf{x}\left(\mathbf{t}_{-i}\right)\right)$ and the bracket follow from (8), so:

$$
\begin{aligned}
t_i\overline{x}_i &\geq \frac{1}{a\left(b+c\right)}\left[\left(\frac{e + \sum_{j\neq i} t_j}{n+1}\right)^2 - cavX_i\right] \\
&\quad - \frac{1}{a\left(b+c\right)}\left[\left(\frac{e - nt_i + \sum_{j\neq i} t_j}{n+1}\right)^2 - ca\left(v + t_i\right)X_i\right] \\
&= \frac{t_i}{a\left(b+c\right)}\left[\frac{2n\left(e + \sum_{j\neq i} t_j\right) - n^2 t_i}{\left(n+1\right)^2} + caX_i\right]. \qquad (28)
\end{aligned}
$$

Thus, D's problem becomes to maximize:

$$
\begin{aligned}
& u_D + \sum_{i\in M}\left[u_i^0\left(\mathbf{x}\left(\mathbf{t}\right)\right) - u_i^0\left(\mathbf{x}\left(\mathbf{t}_{-i}\right)\right)\right] = -dx + \sum_{i\in M}\left[x_i t_i - t_i\overline{x}_i\right] \\
= & -dx + \sum_{i\in M} x_i t_i - \sum_{i\in M}\frac{t_i}{a\left(b+c\right)}\left[\frac{2n\left(e + \sum_{j\neq i} t_j\right) - n^2 t_i}{\left(n+1\right)^2} + caX_i\right].
\end{aligned}
$$

Since x_i is given by (13) and x by (24), the f.o.c. w.r.t. t_i becomes:

$$\begin{aligned}
0 &= \frac{d}{a(b+c)(n+1)} + \frac{(b+c)\overline{p} + ac(n+1)X_i - acX - v - 2t_i n + \sum_{j\neq i} t_j}{a(b+c)(n+1)} \\
&\quad + \frac{1}{a(b+c)(n+1)} \sum_{j\in M\setminus i} t_j - \frac{1}{a(b+c)} \left[\frac{2n\left(e + \sum_{j\neq i} t_j\right) - 2n^2 t_i}{(n+1)^2} + caX_i \right] \\
&\quad - \sum_{j\in M\setminus i} \frac{t_j}{a(b+c)} \frac{2n}{(n+1)^2}.
\end{aligned}$$

Note that X_i disappears from the f.o.c., so we get the same $t_i = t_{PC}$ for every $i \in M$. The f.o.c. thus simplifies to:

$$\begin{aligned}
0 &= (n+1)d + (n+1)e - t_{PC}(2n - m + 1)(n+1) \\
&\quad + (m-1)(n+1)t_{PC} - [2ne - 2n(n-m+1)t_{PC}] - 2n(m-1)t_{PC} \\
&= (n+1)d - (n-1)e - 2[(n-m+1) + n(m-1)]t_{PC},
\end{aligned}$$

which reveals that the second-order condition clearly holds. By solving for t, we get:

$$t_{PC} = \frac{(n+1)d - (n-1)e}{2[(n-m+1) + n(m-1)]} = \frac{(n+1)d - (n-1)e}{2 + 2m(n-1)}.$$

We can find $\overline{x}_i$ by inserting t_{PC} and $x_{i,0}$ from (13) into (28):

$$\begin{aligned}
\overline{x}_i^{PC} &= x_i(\mathbf{0}) + \frac{1}{a(b+c)} \left[\frac{n-1}{(n+1)^2} e + \frac{2n(m-1) - n^2}{(n+1)^2} t \right] \\
&= x_i(\mathbf{0}) + \frac{1}{a(b+c)(n+1)^2} \left[(n-1)e + \left(2n(m-1) - n^2\right) t \right].
\end{aligned}$$

(iii) Note that (IC) is harder to satisfy than (PC) if $\overline{x}_i^{IC} > \overline{x}_i^{PC}$. A simple comparison gives (20). □

Proof of Proposition 6.

(i) It is easy to see that (IC_i) remains unchanged if D can use more general contracts: If D wants to implement a particular vector $\mathbf{x}$, it must offer each $i \in M$ a transfer $\boldsymbol{\tau}_i(\mathbf{x})$ that makes i weakly better off compared to selecting any other x_i (leading to a different transfer). To discourage such deviations, D should ensure that i receives no transfer if i deviates from the implemented plan. Thus, the incentive constraint is

$$u_i^0(\mathbf{x}(\boldsymbol{\tau})) + \boldsymbol{\tau}_i(\mathbf{x}(\boldsymbol{\tau})) \geq u_i^0(\widehat{x}_i, x_{-i}(\boldsymbol{\tau}))\, \forall \widehat{x}_i > \overline{x}_i,$$

just as before.

(ii) Next, note that the participation constraint can always be weakened to make it weaker than the incentive constraint. To see this, write the participation constraint as:

$$u_i^0(\mathbf{x}(\boldsymbol{\tau})) + \boldsymbol{\tau}_i(\mathbf{x}(\boldsymbol{\tau})) \geq u_i^0(\mathbf{x}(\boldsymbol{\tau}_{-i})),$$

and note that it is always possible to select $\boldsymbol{\tau}(\mathbf{x})$ in such a way that $\mathbf{x}_{-i}(\boldsymbol{\tau}_{-i}) = x_{-i}(\boldsymbol{\tau})$, that is, such that no $j \neq i$ will change x_j if i announces that i will not accept transfers from D. This is achieved, for example, if j receives transfers only when $x_j = x_j(\boldsymbol{\tau})$.

Of course, it may be that the transfer $\boldsymbol{\tau}_j$ must be larger when i rejects the contract and thus selects $x_i \neq x_i(\boldsymbol{\tau})$, but this larger transfer will not have to be paid by D in equilibrium.
(iii) Thus, only the incentive constraint will bind when $\boldsymbol{\tau}(\mathbf{x})$ can be a general function. Inserting the binding incentive constraints into D's objective function gives, as before, that D selects $\boldsymbol{\tau}$ or, equivalently, $\mathbf{x}$, to maximize:

$$u_D + \sum_{i \in M} u_i^0(\mathbf{x}) - u_i^0(\widetilde{x}_i(x_{-i}), x_{-i}),$$

where $\widetilde{x}_i(x_{-i}) = \arg\max_{x_i} u_i^0(x_i, x_{-i})$. This is the same problem as in the proof of Proposition 5(i), and the outcome for x_i and $\boldsymbol{\tau}$ are thus also identical. □

Proof of Proposition 7.
From now on we frequently use $y \equiv e/d$. The following proof is more general than needed, since we allow for a regime change that changes $q \equiv n - m$ as well as m (to q' and m'), even though the text above does not consider changes in q. When inserting (19) into (10), we get:

$$x = \frac{c}{b+c}X + \frac{e}{a(b+c)} - \frac{e + 2dm/(n+1)}{a(b+c)(n+1)}.$$

Thus, with decentralization, x increases if:

$$\begin{aligned} 0 &> \frac{e(n'-n)}{(n+1)(n'+1)} - 2d\frac{m'(n+1)^2 - m(n'+1)^2}{(n+1)^2(n'+1)^2} \Leftrightarrow \\ y &< 2\frac{m'(n+1)^2 - m(n'+1)^2}{(n+1)(n'+1)(n'-n)} = \frac{2}{(n'-n)}\left[m'\left(1 - \frac{n'-n}{n'+1}\right) - m\left(1 + \frac{n'-n}{n+1}\right)\right] \\ &= 2\left[\frac{m'-m}{n'-n} - \frac{m'}{n'+1} - \frac{m}{n+1}\right] = 2\left[1 - \frac{q'-q}{n'-n} - \frac{m'}{n'+1} - \frac{m}{n+1}\right]. \end{aligned}$$

Setting $q = q'$ and $\Delta = m' - m = n' - n$ completes the proof. □

Proof of Proposition 8.
With (24) and (27) we can write D's payoff as a function of m:

$$\begin{aligned} -dx - \sum_{i \in M} \frac{t_i^2}{4a(b+c)} &= -d\left(\frac{ac}{a(b+c)}X + \frac{ne - m\left(\frac{2}{n+1}d\right)}{a(b+c)(n+1)}\right) - \frac{m\left(\frac{2}{n+1}d\right)^2}{4a(b+c)} \\ &= -\frac{dcX}{b+c} - \frac{de}{a(b+c)}\frac{n}{n+1} + \frac{2d^2}{a(b+c)}\frac{m}{(n+1)^2} - \frac{m}{(n+1)^2}\frac{d^2}{a(b+c)} \\ &= \frac{d^2}{a(b+c)}\frac{m}{(m+q+1)^2} - \frac{dcX}{b+c} - \frac{de}{a(b+c)}\frac{m+q}{m+q+1}. \end{aligned}$$

By comparison, $m' > m$ increases D's payoff if

$$\begin{aligned}
\frac{m'}{(m'+q+1)^2} - y\frac{m'+q}{m'+q+1} &> \frac{m}{(m+q+1)^2} - y\frac{m+q}{m+q+1} \Leftrightarrow \\
-y\left(\frac{m'-m}{(m+q+1)(m'+q+1)}\right) &> \frac{m}{(m+q+1)^2} - \frac{m'}{(m'+q+1)^2} \Leftrightarrow \\
y(m'-m) &< m' - \frac{m'(m'-m)}{(m'+q+1)} - m - \frac{m(m'-m)}{(m+q+1)} \Leftrightarrow \\
y &< 1 - \frac{m'}{(m'+q+1)} - \frac{m}{(m+q+1)}. \quad \square
\end{aligned}$$

Proof of Proposition 9.
We first derive the equilibrium payoff for a single district. From (26), (8), and (13), we have:

$$\begin{aligned}
u_i^0(\widehat{x}_i, x_{-i}(\mathbf{t})) &= u_i^0(\mathbf{x}(\mathbf{t})) + \frac{t_i^2}{4a(b+c)} \\
&= \left[u_i^0(\mathbf{x}(\mathbf{t})) - x_i t_i\right] + x_i t_i + \frac{t_i^2}{4a(b+c)} \\
&= \frac{1}{a(b+c)}\left[\left(\frac{e - nt_i + \sum_{j\neq i} t_j}{n+1}\right)^2 - ca(v+t_i)X_i\right] + x_i t_i + \frac{t_i^2}{4a(b+c)} \\
&= \frac{1}{a(b+c)}\left[\left(\frac{e-(n-m+1)t}{n+1}\right)^2 - ca(v+t)X_i\right] + \frac{t^2}{4a(b+c)} \\
&\quad + t\frac{(b+c)\overline{p} + ac(n+1)X_i - acX - v}{a(b+c)(n+1)} - t\frac{2d(n-m+1)}{a(b+c)(n+1)^2}.
\end{aligned}$$

With (19), $u_i^0(\widehat{x}_i, x_{-i}(\mathbf{t}))$ becomes

$$\begin{aligned}
&\frac{1}{a(b+c)}\left[\left(\frac{e-(n-m+1)\frac{2}{n+1}d}{n+1}\right)^2 - ca\left(v + \frac{2}{n+1}d\right)X_i\right] + \frac{\left(\frac{2}{n+1}d\right)^2}{4a(b+c)} \\
&+ \frac{2}{n+1}d\frac{(b+c)\overline{p} + ac(n+1)X_i - acX - v}{a(b+c)(n+1)} - \frac{2}{n+1}d\frac{2d(n-m+1)}{a(b+c)(n+1)^2}
\end{aligned}$$

$$\begin{aligned}
= &\frac{1}{a(b+c)}\left[\left(\frac{e-(n-m+1)\frac{2}{n+1}d}{n+1}\right)^2 - ca\left(v + \frac{2}{n+1}d\right)X_i\right] + \frac{d^2}{a(n+1)^2(b+c)} \\
&+ 2d\frac{e + ac(n+1)X_i}{a(b+c)(n+1)^2} - \frac{4d^2(n-m+1)}{a(b+c)(n+1)^3}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{a(b+c)}\left[\left(\frac{e-(n-m+1)\frac{2}{n+1}d}{n+1}\right)^2 - cavX_i\right] + \frac{d^2}{a(n+1)^2(b+c)} \\
&\quad +2d\frac{e}{a(b+c)(n+1)^2} - \frac{4d^2(n-m+1)}{a(b+c)(n+1)^3}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{a(b+c)(n+1)^2}\left[e^2 - 4e\left(\frac{n-m+1}{n+1}d\right) + 4\left(\frac{n-m+1}{n+1}d\right)^2\right] \\
&\quad -\frac{1}{a(b+c)(n+1)^2}\left[-cavX_i(n+1)^2 + d^2 + 2de - \frac{4d^2(n-m+1)}{a(b+c)(n+1)}\right] \\
&= \frac{1}{a(b+c)}\left[\frac{(e+d)^2}{(n+1)^2} - cavX_i - \frac{q+1}{(n+1)^3}4d(e+d) + \frac{4d^2(q+1)^2}{(n+1)^4}\right] \\
&= \frac{1}{a(b+c)}\left[\frac{[(e+d)(n+1)-2d(q+1)]^2}{(n+1)^4} - cavX_i\right] \\
&= \frac{1}{a(b+c)}\left[\left(\frac{(e+d)(n+1)-2d(q+1)}{(n+1)^2}\right)^2 - cavX_i\right].
\end{aligned}
$$

Consider now a set L of $l = |L|$ districts, taking as given $n-l$. The sum of L's payoffs is:

$$
\sum_{i\in L}\frac{1}{a(b+c)}\left[\left(\frac{(e+d)(n+1)-2d(q+1)}{(n+1)^2}\right)^2 - cavX_i\right].
$$

If the set L decentralizes, the new number is larger, say $l' = l + \Delta > l$. This reduces total welfare if:

$$
\begin{aligned}
\frac{l}{a(b+c)}\left(\frac{(e+d)(n+1)-2d(q+1)}{(n+1)^2}\right)^2 &> \frac{l'}{a(b+c)}\left(\frac{(e+d)(n'+1)-2d(q+1)}{(n'+1)^2}\right)^2 \Leftrightarrow \\
\frac{l}{l'}\left(\frac{n'+1}{n+1}\right)^2 &> \left(\frac{e+d\frac{n'-2q-1}{n'+1}}{e+d\frac{n-2q-1}{n+1}}\right)^2 = \left(\frac{(y+1)-\frac{2(q+1)}{(n'+1)}}{(y+1)-\frac{2(q+1)}{(n+1)}}\right)^2 \Leftrightarrow \\
\frac{l}{l'}\left(\frac{n'+1}{n+1}\right)^2 &> \Omega^2, \text{ where } \Omega \equiv 1 + \frac{\frac{2m'}{n'+1}-\frac{2m}{n+1}}{y+\frac{2m}{n+1}-1}. \qquad (29)
\end{aligned}
$$

The r.h.s. of (29), Ω^2, is, as a function of $y = e/d$, drawn in the figure. When $y \in \left(-\infty, 1-\frac{2m'}{n'+1}\right)$, $\Omega > 0$ and Ω^2 decreases from 1 to 0. When $y \in \left(1-\frac{2m'}{n'+1}, 1-\frac{2m}{n+1}\right)$, $\Omega < 0$ and Ω^2 increases from 0 to ∞. When $y > 1-\frac{2m}{n+1}$, $\Omega > 0$ and Ω^2 decreases from ∞ to 1 when y increases toward ∞.

Thus, when $y \in \left(1-\frac{2m'}{n'+1}, 1-\frac{2m}{n+1}\right)$ and $\Omega < 0$, the inequality (29) can be written

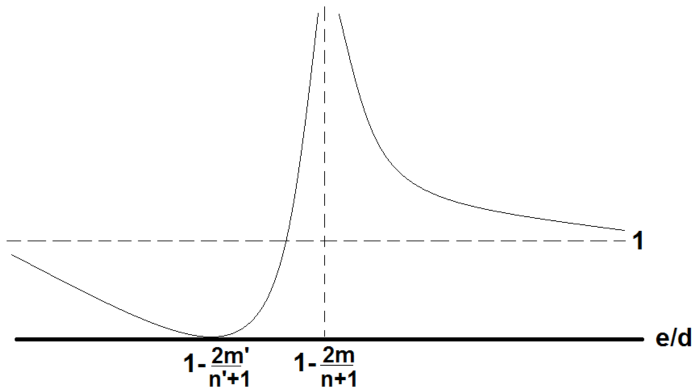

Figure 3: Ω^2 *as a function of e/d*

as:

$$\sqrt{\frac{l}{l'}}\left(\frac{n'+1}{n+1}\right) > \frac{\frac{2m'}{n'+1}-\frac{2m}{n+1}}{1-y-\frac{2m}{n+1}}-1 \Leftrightarrow$$
$$y < \widehat{\epsilon}_L \equiv 1-\frac{2m}{n+1}-\frac{\frac{2m'}{n'+1}-\frac{2m}{n+1}}{\sqrt{\frac{l}{l'}}\left(\frac{n'+1}{n+1}\right)+1},$$

which clearly satisfies $y \in \left(1-\frac{2m'}{n'+1}, 1-\frac{2m}{n+1}\right)$. In the example with $m = n = 2$, we get $\widehat{\epsilon}_L \approx -0.16$.

If instead $\Omega > 0$, (29) implies:

$$\sqrt{\frac{l}{l'}}\left(\frac{n'+1}{n+1}\right) > 1+\frac{\frac{2m'}{n'+1}-\frac{2m}{n+1}}{y+\frac{2m}{n+1}-1}. \tag{30}$$

If, moreover, the denumerators are positive, then (30) implies:

$$y > \bar{\epsilon}_L \equiv 1-\frac{2m}{n+1}+\frac{\frac{2m'}{n'+1}-\frac{2m}{n+1}}{\sqrt{\frac{l}{l'}}\left(\frac{n'+1}{n+1}\right)-1}.$$

In our example, $\bar{\epsilon}_L = \frac{4}{9\sqrt{2}-12} \approx 5.5$.

If instead the denumerator on the r.h.s. of (30) is negative, so that $y < 1-\frac{2m'}{n'+1}$, then (30) implies:

$$y < \bar{\epsilon}_L \equiv 1-\frac{2m}{n+1}+\frac{\frac{2m'}{n'+1}-\frac{2m}{n+1}}{\sqrt{\frac{l}{l'}}\left(\frac{n'+1}{n+1}\right)-1}.$$

Note also that $\Omega < -1$ (which implies $\Omega^2 > 1$) if:

$$y \leq 1-\frac{m}{n+1}+\frac{m'}{n'+1}.$$

Finally, note the following:

Lemma: *The l.h.s. of (29) is larger than 1 if and only if L is large:*

$$\frac{l}{l'}\left(\frac{n'+1}{n+1}\right)^2 > 1 \Leftrightarrow \frac{l'}{1+n'} + \frac{l}{n+1} > 1.$$

Proof: Note that

$$\begin{aligned}
\frac{l}{l'}\left(\frac{n'+1}{n+1}\right)^2 &> 1 \Leftrightarrow && (31)\\
\frac{l'-v}{l'}\frac{(n'+1)^2}{(n'+1)^2-2v(n'+1)+v^2} &> 1 \Leftrightarrow &&\\
\frac{v}{l'} &< \frac{2v(n'+1)-v^2}{(n'+1)^2} \Leftrightarrow &&\\
\frac{l'}{1+n'} &> \frac{n'+1}{2(n'+1)-v}. && (32)
\end{aligned}$$

In addition, (31) is equivalent to

$$\begin{aligned}
\frac{2(n+1)v+v^2}{(n+1)^2} &> \frac{v}{l} \Leftrightarrow &&\\
\frac{l}{n+1} &> \frac{n+1}{2(n+1)+v}. && (33)
\end{aligned}$$

Since (32) and (33) are equivalent, we can also sum them and write:

$$\begin{aligned}
\frac{l'}{1+n'} + \frac{l}{n+1} &> \frac{n'+1}{2(n'+1)-v} + \frac{n+1}{2(n+1)+v}\\
&= \frac{n'+1}{n+n'+2} + \frac{n+1}{n+n'+2} = 1. \qquad \square
\end{aligned}$$

Proof of Proposition 10.
(i) If the presence of D leads to centralization and $n' > n$, then x increases if:

$$\begin{aligned}
\frac{e}{a(b+c)(n+1)} &> \frac{e+2dm'/(n'+1)}{a(b+c)(n'+1)} \Leftrightarrow\\
y(n'-n) &> \frac{2m'(n+1)}{n'+1} \Leftrightarrow\\
y &> \frac{2m'(n'+1-v)}{(n'+1)v} = \frac{2m'(n+1)}{(n'+1)v} = 2\left(1+\frac{m}{v}\right)\left(1-\frac{v}{n'+1}\right)\\
&= \frac{2(m+v)(n+1)}{(n+v+1)v}.
\end{aligned}$$

In the example, the condition becomes $y > 8/3$.

(ii) If the presence of D leads to centralization and thus $n' < n$, then x increases if

$$\begin{aligned}
\frac{e}{a(b+c)(n+1)} &> \frac{e+2dm'/(n'+1)}{a(b+c)(n'+1)} \Leftrightarrow \\
y(n'-n) &> \frac{2m'(n+1)}{n'+1} \Leftrightarrow \\
y &< -\frac{2m'(n+1)}{(n'+1)v}. \quad \square
\end{aligned}$$

Proof of Proposition 11. The proof follows the same steps as above, except that it is longer, since we cannot aggregate the terms when all parameters are asymmetric. We have thus left out the proof here, in the interest of space, but we will make it available on request. □

Proof of Proposition 12.
A nonparticipant's $i \in L \subset N \backslash M$ payoff is, from (8),

$$\begin{aligned}
u_i &= \frac{1}{a(b+c)}\left[\left(\frac{e+mt}{n+1}\right)^2 - cavX_i\right] = \frac{1}{a(b+c)}\left[\left(\frac{e+m2d/(n+1)}{n+1}\right)^2 - cavX_i\right] \\
&= \frac{1}{a(b+c)}\left[\left(\frac{e+2dm/(n+1)}{n+1}\right)^2 - cavX_i\right],
\end{aligned}$$

so the sum for L is, when $n-l$ is given:

$$\sum_{i\in L}\frac{1}{a(b+c)}\left[\left(\frac{e+2dm/(n+1)}{n+1}\right)^2 - cavX_i\right].$$

This sum increases after decentralization (i.e., for $l' > l$) if

$$\begin{aligned}
l'\left(\frac{e+2dm/(n'+1)}{n'+1}\right)^2 &> l\left(\frac{e+2dm/(n+1)}{n+1}\right)^2 \Rightarrow \\
\frac{l}{l'}\left(\frac{n'+1}{n+1}\right)^2 &< \left(\frac{e/d+m/(n'+1)}{e/d+m/(n+1)}\right)^2.
\end{aligned}$$

The rest of the proof follows the analoguous steps as in the proof of Proposition 9. □