Contracts and Induced Institutional Change*

Bård Harstad Torben Mideksa

July 20, 2021

Abstract

We study agents' incentives to form horizontal coalitions before a principal offers vertical contracts. When one contract generates negative externalities on other agents, the agents may collude in order to obtain better deals; when one contract benefits other agents, the agents may decentralize, instead. Contractually induced institutional changes always harm the principal and the negative effect can outweigh the direct effects of the contracts, making the contracts counterproductive. The model is tractable and sufficiently flexible to be relevant for applications such as regulation of pollution, payments for forest conservation, and mergers between firms in a supplier-franchisee relationship.

Keywords: vertical contracts, centralization, decentralization, induced institutional change, conservation, pollution regulation, retailers

JEL codes: D86, H23, H87, Q58

*We thank the editor, Mike Powell, three excellent referees, as well as Daron Acemoglu, Arild Angelsen, Philippe Delacote, Jonas Hjort, David Martimort, Chuck Mason, Halvor Mehlum, Kalle Moene, Nicola Persico, Torsten Persson, Francois Salanie, Steve Shavell, Kathryn Spier, Jon Strand, Jean Tirole, and Ragnar Torvik. The manuscript has also benefited from comments of audiences at Brown University, Harvard University, London School of Economics, NYU, Norwegian School of Economics, Paris Ecole Polytechnique, Stanford GSB, Toulouse School of Economics, UC Berkeley, Universidad Carlos III, University of Copenhagen, University of Oslo, University of Zurich, Warwick University, the World Bank, and the 2015 Environmental Protection and Sustainability Forum in Bath. Frank Azevedo assisted with the editing and Kristen Vamsæter with the figures.
Emails: bard.harstad@econ.uio.no and torben_mideksa@post.harvard.edu.

1 Introduction

In this article, we study how vertical contracts – or the mere anticipation of contracts – can motivate agents to either centralize or decentralize authority.

There are several real-world cases in which it is important to recognize how contracts influence coalitions: (a) if a regulator taxes or subsidizes emission reductions, the firms may anticipate that the policy stringency will have to be modified if the firms merge; (b) if a donor pays for reducing emissions from deforestation and forest degradation (REDD+), the payment levels may have to increase if the recipient districts (de)centralize forest management; (c) if an upstream supplier provides an input to several retailers, the retailers may merge or decentralize to obtain agreements that are more attractive.

To develop an understanding of these cases, we begin by considering a simple and stylized Cournot game, with linear demand, where we permit pecuniary as well as (positive or negative) non-pecuniary externalities. In addition to the agents, we consider a principal that is not indifferent to the firms' production levels. The principal can offer payments, conditional on the quantities, to a subset of the agents.

Consider now the possibility that agents collude in the absence of the principal. If the externalities from reduced production are positive, a set of k firms may collude or merge to reduce production and thus externalities. The other $n-k$ firms in the market will react by raising their production quantities, however, undoing parts of the coalition's effort. In line with the logic of Coasian bargaining, we consider the possibility that a set of agents centralize if and only if centralized decision making is expected to increase the sum of payoffs for this set. Because of the non-members' response, k firms benefit from centralizing only if k is large, whereas they prefer decentralization if k is small (Proposition 0). The importance of the size of k, relative to that of n, is a standard finding in the literature (see the literature review).

Next, we combine the vertical contracts and the horizontal mergers. Because institutional changes take time, the natural timing is (1) a subset of agents decides whether or not to merge/centralize, (2) the principal commits to payments as a function of quantities, before (3) agents simultaneously choose their outputs to maximize payoffs.

The anticipation of the principal's offer can motivate contractually induced institutional change. If, for example, reduced production of a firm benefits both the principal and the other firms, then the principal can pay less if the firms centralize authority and thereby internalize the positive externality from the contract. In this case, therefore, the firms may benefit from decentralizing authority to increase the payments. If reduced production harms other agents, as when the pressure of illegal deforestation increases in other districts, a central authority would require larger payments to conserve. In this case, districts may centralize to obtain more attractive contracts.

Depending on the ratio of the externality between agents to the externality facing the principal, small groups of agents may be induced to centralize, or large groups to decentralize (Proposition 1), reversing the equilibrium arrangement without contracts (i.e., Proposition 0). In our model, a contractually induced institutional change will always reduce the principal's payoff (Proposition 2) and the change will increase the harm facing the principal (Proposition 3). Under specified conditions, the harm can be larger than it would have been if the principal had been able to commit to offer nothing (Proposition 4). Our quantitative assessment shows not only that the induced changes are a theoretic possibility, but also that they arise for a large set of parameters. They can also shed light on all applications (a)–(c), discussed above.

Literature and evidence. Thanks to the breadth of our applications, the article contributes to strands of literature ranging from industrial organization to conservation.

In industrial organization, it is well known that firms may prefer to collude or merge to gain market power (see Marshall and Marx, 2012, for a detailed textbook treatment). However, if a set of firms merges to cut production, firms outside of the merger might decide to increase *their* production levels. When these effects are taken into account, a merger is profitable for k firms only when k is a majority of all firms in the market (Salant et al., 1983). This result has later been extended in various directions (Gaudet and Salant, 1991; Kamien and Zang, 1990; Perry and Porter, 1985). The condition derived by Salant et al. (1983:191) is a special case of the condition we present below for the simplest benchmark case in which there is no principal or contract (Proposition 0).

The basic effects of mergers on prices (and thus quantities) are extensively documented empirically. Hastings (2004) and Houde (2012) have shown that consolidation in the gasoline retail market increases prices when refineries are active players in the vertical relation. Silvia and Taylor (2013) argue that the prices increased when Sunoco acquired El Paso in 2004, and Valero acquired Premcor's Delaware refinery in 2005.[1] Given these empirical findings, it is important to study, as we do, how the incentives to merge change when vertical contracts are anticipated.

Earlier articles have emphasized that downstream firms may want to collude or merge to increase their bargaining power when dealing with upstream suppliers: See Inderst and Shaffer (2007) and Inderst and Wey (2003), who also studied how the merger may influence the supplier's technology choice. Inderst and Wey built on Horn and Wolinsky (1988), who examined the incentives for mergers between duopolies given the effects on input and output prices. Milliou and Petrakis (2007) extended this approach further by allowing for different types of vertical contracts. (See also McAfee and Schwartz, 1994, or Dequiedt and Martimort, 2015, who compare publicly vs. privately observable vertical contracts.)

The above effects, which are present also in our analysis, are empirically plausible: The United Kingdom's Competition and Markets Authority (2019) worry that consolidation in the U.K. grocery retail industry might indeed have weakened suppliers' bargaining power.[2] Barro and Cutler (1997) show that three out of four mergers among hospitals in Massachusetts were motivated by the need to strengthen their bargaining power or leverage relative to that of insurance companies. Since such mergers can reduce the costs for the entities that merge, possibly benefitting consumers, these types of mergers have been defended in courts (Dranove et al., 2019).

Since our "contracting with externalities" model draws on Segal (1999), we contribute to the associated literature, such as the studies by Segal (2003) and Che and Spier (2008) on how a principal may ex ante design a contract that "divides and conquers." (See also Genicot

[1] Focusing on the empirical aspect of mergers, Hosken and Tenn (2015) describe antitrust analyses of horizontal mergers in U.S. retail markets and offer a very rich account of various merger cases such as those of Vons supermarkets, Syufy Enterprises, Staples, and Whole Foods Markets, Inc. Ashenfelter and Hosken (2010) examine five mergers and show that in four of those cases, the mergers raised market prices. Ashenfelter, Hosken, and Weinberg (2013) study the effect of a large manufacturing merger and find that the effect on market prices is substantial. Hosken, Olson, and Smith (2018) study the consequences of grocery mergers in 14 regions, including mergers in highly concentrated and relatively unconcentrated markets. They show that mergers raise prices, especially in highly concentrated markets.

[2] Landeo and Spier (2009) show that, in experiments, even nonbinding communication among the agents (or among the buyers, in their framework) may be sufficient for improving the offer made by the principal. Their paper is related to that of Asker and Bar-Isaac (2014), who study how vertical contracts can incentivize retailers to not accommodate entrant manufacturers.

and Ray, 2006, who study how a principal can get agents to sign contracts when externalities are negative.) A natural question, in these situations, is how agents can benefit from either decentralizing or centralizing before contracts are offered. Our contribution is to shed light on this question.

In the literature on deforestation, Burgess et al. (2019) explain that the deforestation rate can be influenced and reduced by governmental policies even when logging is illegal. Burgess et al. (2012) find that decentralization has increased deforestation in Indonesia, but Agrawal et al. (2008), Irawan and Tacconi (2009), and Berkes (2010) have found that decentralization reduced deforestation in many other countries. The conflicting empirical evidence can be explained by our basic model if deforestation is profit-driven in Indonesia, but driven by illegal logging in the other countries. That explanation is discussed in Harstad and Mideksa (2017). In contrast to our earlier article, we here endogenize the agents' decision on whether to centralize or decentralize, and we study how the agents' preferences regarding the institutional structure depend on the anticipated vertical contracts. Our present model is also more general in several ways. (Our earlier article focused on the application (b) about deforestation.)[3]

Outline. In the next section, we present the basic model before we discuss the three interpretations. In Section 3, we first describe when centralization is beneficial for a subset of agents in the absence of the principal, before we show when, and how, the anticipation of vertical contracts can induce institutional changes. Section 3 also shows that an induced institutional change will always harm the principal, and we describe when the overall effect on quantities and harm outweighs the direct effects of the contracts. In section 4 we discuss numerical assessments of the predictions and the implications for the three applications. Section 5 concludes and the Appendix contains all proofs.

[3]Since we here focus on the harmful effects of induced institutional change, we contribute to the literature on explanations of deforestation. See, for example, Burgess et al. (2012), Amacher et al. (2012), Amacher et al. (2007), Mendelsohn (1994), Angelsen and Kaimowitz (1999), and the references therein. Furthermore, our focus on institutional change is different from the focus in the papers on REDD+, which emphasize moral hazard (Gjertsen et al., 2020), private information (Chiroleu-Assouline et al., 2012; Mason, 2015; Mason and Plantinga, 2013), or observability (Delacote and Simonet, 2013).

2 Markets and contracts

2.1 Agents and markets

In this section, we present a simple principal–agent model before we discuss four interpretations of it.

Agents. The basic model is essentially a classic Cournot game, extended to allow for non-pecuniary externalities as well as contracts. There are n producers or "agents," simultaneously taking action $x_i \geq 0$, $i \in N \equiv \{1, ..., n\}$. Agent i ("he") produces $x_i \geq 0$ units at marginal cost $c > 0$ and sells them in the common market at price p:

$$p = \overline{p} - ax_N \text{ and } x_N \equiv \sum_{i \in N} x_i. \tag{1}$$

The common market means that there are pecuniary externalities between the firms. Externalities are central in this paper and for some of the applications discussed in Section 2.3, such as when there is pollution, it is natural to include non-pecuniary externalities or spillovers (s_i). Agent i's payoff is given by:

$$u_i(x_i, x_{-i}) = wpx_i - cx_i - s_i x_N + t_i. \tag{2}$$

Parameter $w > 0$ measures the weight of the gross revenue (px_i) relative to direct transfer from contracting (t_i). It is, for many applications, natural to simply set $w = 1$. Application (b) on deforestation requires $w \neq 1$, however.

The market. If we abstract from the transfers ($t_i = 0 \forall i \in N$), the agents play a simple Cournot game. The first-order condition of (2) is satisfied when:

$$x_i = \frac{w\overline{p} - wax_{N \setminus i} - c - s_i}{2wa}. \tag{3}$$

As in any Cournot model with quantity competition, actions are strategic substitutes, so i produces less when the others produce more.

If we combine the n first-order conditions, we find that the unique Nash equilibrium is:

$$x_i^0 = \frac{w\bar{p} - c + s_N - (n+1)\, s_i}{wa\,(n+1)}, \text{ where } s_N \equiv \sum_{i \in N} s_i, \text{ so} \tag{4}$$

$$x_N^0 = \frac{nw\bar{p} - nc - s_N}{wa\,(n+1)}. \tag{5}$$

The horizontal externalities between the agents will be important when we discuss coalition formation. The envelope theorem implies that j's decision to reduce the amount of activity x_j has the following effect on u_i:

$$\begin{aligned} \frac{\partial u_i\left(x_i, x_{N\setminus i}\right)}{\partial\,(-x_j)} &= awx_i^0 + s_i = \frac{\varepsilon}{n+1}, \text{ where} \\ \varepsilon &\equiv w\bar{p} - c + s_N. \end{aligned} \tag{6}$$

Note that $\varepsilon > 0$ implies net negative externalities of production, so agents experience positive net externalities when another agent produces *less*.

The above equations are sensible only when $p \geq 0$ and $x_i^0 \geq 0 \forall i \in N$. These inequalities hold if and only if:

$$\varepsilon \geq \max_i \left\{-(n+1)\, c, (n+1)\, s_i\right\}, \tag{7}$$

which we will assume.

The comparative statics are intuitive and standard. A producer with a lower total marginal cost, $c + s_i$, relative to that of the others, $\sum_{j \in N \setminus i} (c + s_j)$, produces more. A smaller a or a larger $\bar{p}$ increases the price and the equilibrium quantities that are produced. The object ε is the combined externality from agent j's reduction in x_i on the payoff of agent i. The externality is large if the weight on profit w is large, if the market size $\bar{p}$ is large, and if s_N is large. In such a situation, i benefits a lot when the other producers reduce the supply since that reduction raises the market price and i's profit.

2.2 Principal and contracts

The principal. The direct transfer to i comes from the principal (P, "she"). The alternative roles of P are given in the next subsection. We assume that P can contract with the set $M \subseteq N$

of "members." The reader is free to restrict attention to the case in which $M = N$, so that P contracts with all the agents. Allowing for $M \subseteq N$ does not complicate the analysis, however, and it generates additional results. P may benefit from adjustments in the production levels of the $m \equiv |M|$ members as well as from adjustments in the production levels of those agents that are left without contracts, $L \equiv N \backslash M$, where $l \equiv |L|$. To keep it all simple, let P's objective function be:

$$u_P = -d_M x_M - d_L x_L - t_M, \tag{8}$$

where the subscript denotes the set of agents we aggregate over, so that $t_M \equiv \sum_{i \in M} t_i$, $x_L \equiv \sum_{i \in L} x_i$, etc. Parameter d_M represents the marginal damage P experiences from the production of agents with whom she can contract, and d_L measures the damage from the production of the others. So, in addition to the costly transfers, we can write the total harm experienced by P as:

$$h \equiv d_M x_M + d_L x_L.$$

Whether d_M, d_L, and h are positive or negative will depend on the application at hand. If there are no contracts, we can use (4) to show that the harm is:

$$h^0 = \frac{d_M}{aw} \left(\frac{m\varepsilon}{n+1} - s_M \right) + \frac{d_L}{aw} \left(\frac{l\varepsilon}{n+1} - s_L \right).$$

Contracts. Before the agents set their x_i, which are observable and verifiable, the principal commits to positive transfer schedules. The transfers are paid to every $i \in M$ at the end of the game as a function of $\mathbf{x} = (x_1, ..., x_n)$. The transfer to $i \in M$, $t_i(\mathbf{x}) \geq 0$, can be a general function of $\mathbf{x} = (x_1, ..., x_n)$. The transfer schedules are verifiable and observable to everyone. The contracts are characterized by *limited liability* in that $i \in M$ is free to ignore the transfer schedule and select any x_i he wants. Thus, the contract is, in the model, binding only for P.

Note, however, that it would not make any difference if we introduced an acceptance stage in which every $i \in M$ could approve or decline the transfer schedule proposed by P. If P wants to implement a vector $\mathbf{x}^*$, P can discourage any up-front rejection by designing $\mathbf{t}(\mathbf{x}) = (t_1(\mathbf{x}), t_2(\mathbf{x}), ...)$ so that if an agent i declines, then the schedule $\mathbf{t}(\mathbf{x}_{-i}; x_i \neq x_i^*)$ will motivate the other members to select a vector $\mathbf{x}_{-i}$ that is very harmful to i. In this way, P can implement any $\mathbf{x}^*$ she wants, at no cost, if it were not for the limited liability constraint that we introduced in the previous paragraph.

Based on this reasoning, we can assume away the possibility that $i \in M$ can up-front reject the transfer schedule proposed by P. With that assumption, it also follows that if a set of functions, $\mathbf{t}(\mathbf{x})$, implement $\mathbf{x}^*$, then $\mathbf{x}^*$ will also be an equilibrium if the principal simply offers the fixed payment $t_i = \mathbf{t}_i(\mathbf{x}^*) \geq 0$ to $i \in M$ when $\mathbf{x} = \mathbf{x}^*$, and zero otherwise. When P can select the equilibrium, we can without loss of generality restrict attention to this type of transfer schedule. In this case, agent $i \in M$ is free to decide on any level x_i, but i recognizes that there will be no transfer if $x_i \neq x_i^*$.

With all the above, the problem for the principal is:

$$\max_{\{t_i, x_i^*\}_{i=1}^m} \{-d_M x_M^* - d_L x_L - t_M\},$$

subject to the following m incentive constraints,

$$u_i\left(x_i^*, x_{N\setminus i}^*\right) + t_i \geq \max_{\widehat{x}_i \geq 0} u_i\left(\widehat{x}_i, x_{-i}^*\right), \tag{IC}$$

and subject to the constraint that x_L depends on x_M^*. From (3), we get:

$$x_L = \frac{l}{1+l}\frac{w\bar{p} - wax_M - c - s_i}{wa} \Rightarrow \frac{\partial x_L}{\partial x_M} = -\frac{l}{1+l} < 0.$$

Before presenting the solution to the principal's problem, it is worth studying a simple condition for when a reduction in x_M reduces the harm, h:

$$\begin{aligned} \frac{\partial h}{\partial x_M} &= d_M \cdot \left(\frac{\partial x_M}{\partial x_M}\right) + d_L \cdot \left(\frac{\partial x_L}{\partial x_M}\right) = \frac{\mu}{1+l}, \text{ where} \\ \mu &\equiv d_M + l\left(d_M - d_L\right). \end{aligned} \tag{9}$$

The vertical externality μ will be essential in our analysis, because the total harm increases in the contracted-upon x_M if and only if $\mu > 0$. It is clearly possible that $\mu < 0$ even if $d_M > 0$, if just d_L is much larger than d_M. For this preference to hold, it is necessary, but not sufficient, that $d_L > d_M$. In addition, we must have that $d_L - d_M > d_M/(l)$.[4]

If we solve the problem of the principal, subject to the above constraints, we can characterize the unique subgame-perfect equilibrium of the game.

Lemma 1. *In equilibrium, P pays t_i^* to $i \in M$ when $x_i = x_i^*$, where:*

$$\begin{aligned} t_i^* &= \frac{1}{aw}\left(\frac{\mu}{n+1}\right)^2 \text{ and} \\ x_i^* &= x_i^0 - 2\mu\frac{l+1}{aw(n+1)^2}, \text{ implying} \\ h &= h^0 - \frac{2m}{aw}\left(\frac{\mu}{n+1}\right)^2. \end{aligned} \tag{10}$$

As in Section 2.1, the equations are sensible only if they give $p \geq 0$ and $x_i^* \geq 0 \forall i \in N$. These inequalities require:

$$\varepsilon \geq \max_i \left\{ (1+n)s_i + 2\frac{1+l}{1+n}\mu, (1+n)s_i - \frac{2m}{1+n}\mu, -(1+n)c - \frac{2m}{(1+n)}\mu \right\},$$

which coincides with (7) if $\mu = 0$.

The comparative statics are intuitive. When $\mu > 0$ is larger, the principal pays more to encourage a smaller x_i^* for $i \in M$. Even though the non-members produce more, there is always a reduction in total harm. If $\mu < 0$, the principal pays $i \in M$ to increase production, $j \in L$ produces less, and, also in this case, h is reduced relative to the situation with no contracts.

P cares about the x_i's but she also desires to reduce the total payments. This desire implies that a larger m reduces the contractually induced change in x_i^*, holding μ constant (note that μ does not vary with m if $d_M = d_L$). The intuition is that when P pays for a change in x_i, then it becomes more expensive to encourage every other $j \in M \backslash i$ to alter x_j in the same direction, given that the x_i's are strategic substitutes. Therefore, when P contracts with a larger number of agents, the change in each x_i is smaller, although the reduction in h is, nevertheless, larger.

Furthermore, the transfer to $i \in M$ is smaller when n is large. The reason is that when n is large, the "leakage" is large, as well, and this leakage discourages P from spending a lot of

[4] As discussed in Section 2.3, $\mu < 0$ is possible in the application (a) on pollution when P regulates gas producers competing with coal producers. In this case, it is possible that the regulator prefers to subsidize gas production since gas replaces coal.

funds for an adjustment in x_i.[5]

2.3 Interpretations and applications

The purpose of this subsection is to illustrate that even though the model is simple and stylized, it is sufficiently flexible to capture a wide range of important applications. We will return to the various applications after the analysis.

(a) Pollution. The market can be for any type of good, if just the production of x_i is associated with some pollution. The principal, then, may be a domestic regulator that internalizes the social damage associated with pollution, and that seeks to reduce pollution using the subsidies $\mathbf{t}(\mathbf{x})$. The social marginal damage can, naturally, be different for the regulated firms (d_M) than for other firms (d_L).

Alternatively, the agents can be countries extracting fossil fuels, whereas P is a climate coalition paying countries to extract less (as in Harstad, 2012; Collier and Venables, 2014; or Steckel, Edenhofer, and Jakob, 2015). In this case, direct spillovers between the producer countries are quite natural, because of climate change, and thus i benefits when $j \neq i$ conserves (i.e., $s_i \geq 0$ and $\varepsilon > 0$). In line with the model, it is for this application reasonable that P can subsidize, but not tax, the agents (i.e., $t_i \geq 0$), and that an agent cannot be punished (beyond $t_i = 0$) no matter what quantity is selected.

With different types of fossil fuels, $d_M \geq 0$ can measure the CO_2 emissions when $i \in M$ produces a unit of energy, while $d_L \geq 0$ can measure the CO_2 emissions when $i \in L$ produces energy. If the regulator, P, pays coal producers to conserve, while the competitors supply natural gas, then, because the CO_2 content is larger in coal than in gas, $d_M > d_L > 0$. From (9), $\mu > 0$.

Alternatively, when a single country, such as Norway, regulates its domestic natural gas

[5] In this model, it turns out to be sufficient for the principal to consider linear contracts of the type $t_i(\mathbf{x}) = \max\{0, (\overline{x}_i - x_i)\tau_i\}$, where $\overline{x}_i$ is the baseline for i, and $\tau_i \geq 0$ is the subsidy per reduced unit of activity. When $d_M = d_L = d$, the outcome in Lemma 1 can be implemented by the linear contract:

$$\tau_i = \frac{2d}{n+1}, \text{ and } \overline{x}_i = x_i^0 + \frac{4m - 3(n+1)}{4aw(n+1)}\tau_i,$$

where x_i^0 is i's production level in the absence of any contracts. For a fixed n, a larger m means that the benchmark $\overline{x}_i$ will be larger, and thus the transfer to i will be larger (for any $x_i < \overline{x}_i$ that i may choose). Norway's existing REDD+ contracts are indeed taking this form.

producers competing with foreign coal producers, then $d_L > d_M > 0$. If l is large, (9) implies that $\mu < 0$. In this case, Lemma 1 implies that P may want to subsidize domestic production. The intuition is that such subsidies can crowd out dirtier foreign production.[6]

(b) Deforestation and REDD+. In our second application, the agents are different jurisdictions, or districts, endowed with tropical forests.[7] Suppose that district $i \in N$ is endowed with a forest of size X_i. A part, $x_i \in [0, X_i]$, is logged, perhaps illegally. The part that is not logged, $X_i - x_i$, must be protected, requiring that on each unit of that forest the monitoring must be so frequent that the expected penalty outweighs the profit from logging – as measured by the market price for timber, p. Let $\kappa > 0$ represent the additional monitoring cost that is necessary if a district seeks to raise the expected penalty, by a marginal unit, on one unit of the forest. Then, κp is the cost of discouraging illegal logging on each unit. District i's total monitoring cost is $(X_i - x_i)\,\kappa p$.

Furthermore, let $b > 0$ measure the districts' weight on the revenues, px_i. Even when logging is illegal, the weight b can measure the fraction of the revenues that are captured by the police at the border. Alternatively, b can capture how much i cares about the revenues consumed by poor farmers. Here, $c > 0$ can measure i's marginal environmental (or opportunity) cost of deforestation. Then, i's payoff can be written as:

$$u_i = bpx_i - \kappa p\,(X_i - x_i) - cx_i + t_i = wpx_i - s_i x_N - cx_i + t_i,$$

just as in (2), minus the constant $\kappa\overline{p}X_i$, when we define

$$s_i \equiv -\kappa a X_i < 0 \text{ and } w \equiv b + \kappa. \tag{11}$$

For this application, $s_i < 0$, and thus, possibly, $\varepsilon < 0$. The intuition is that more conservation in district $j \in N$ increases p and, with that, the monitoring cost for district $i \in N \backslash j$.

Note that (11) necessitates heterogeneous s_i's when the forest size varies among the districts. Interestingly, the deforestation application endogenizes and captures the intuition that a larger

[6]For these applications, note that the model of the market, in Section 2.1, simplifies by assuming that coal and natural gas are perfect substitutes.

[7]For further motivation of this model, see Harstad and Mideksa (2017).

(or merged) producer, with a larger X_i, will produce at a lower marginal cost $(c - -\kappa a X_i)$. Here, this intuitive result follows because when i's forest is large, it becomes more important to reduce the monitoring cost, which is achieved by increasing x_i.

To reduce the damage from deforestation, Norway and the World Bank contribute to the United Nations Programme for Reducing Emissions from Deforestation and Forest Degradation (REDD+). The REDD+ payments, offered in return for reduced deforestation, are similar to the payment schedules described in Section 2.2. (See, in particular, the footnote after Lemma 1.)

For this application, it makes sense that a donor (such as the World Bank, an NGO, or a national government such as Norway's) has the utility function assumed in (8). If the environmental damage from deforestation is the same, regardless of where deforestation takes place, then $d_M = d_L > 0$, implying $\mu > 0$. We have $d_M > d_L > 0$ and, again, $\mu > 0$, if the higher damage from deforestation in M is the reason for why P pays exactly these countries to conserve.

(c) Supplier and retailers. Other applications are unrelated to the environment. In fact, N can be the set of producers selling any type of good. If there is no direct spillover, $\varepsilon > 0$. The principal can be an upstream firm, supplying a valuable input to $i \in M \subseteq N$. In particular, suppose that P's input can reduce the marginal production cost from c to $\underline{c} < c$. In this case, P requires a transfer, $T_i(x_i)$, from i to P in return for the franchise, or for the inputs needed to produce an agreed-upon quantity, x_i. This transfer enables P to capture the benefit of i's lower production costs. If we define $t_i \equiv (c - \underline{c}) x_i - T_i(x_i)$, then i's objective function can be written as (2), as before. The total revenue to the supplier becomes:

$$u_P = \sum_{i \in M} T_i(x_i) = (c - \underline{c}) \sum_{i \in M} x_i - t_M = -d_M x_M - d_L x_L - t_M,$$

just as in (8), when we define:

$$d_M = -(c - \underline{c}) < 0 \text{ and } d_L = 0.$$

Thus, for the supplier-and-retailers application, $d_M < 0$ and $d_L = 0$ imply that $\mu < 0$. The negative μ implies that, fixing the t_i's, P benefits if $i \in M$ produces and purchases more of P's input.[8]

3 Institutional change

We now introduce the final building block of our model, namely how a set of agents may decide whether to (de)centralize. Section 3.2 combines the three parts of the model to show when we may expect induced institutional change. In the following three subsections, we study the consequences of induced institutional change for the principal and for the damages.

3.1 Institutions without contracts

The horizontal externality in (6) implies that the decentralized outcome is not Pareto efficient for the agents, and that the sum of payoffs can be larger if the production decisions are coordinated, or centralized. If the externality ε is positive, a reduction in the production level of one agent raises other agents' payoffs. If the externality is negative (e.g., as in the application (b) on deforestation), a reduction in x_i reduces other agents' payoffs. In this subsection, we analyze how these horizontal externalities can influence institutional changes.

Consider a coalition $K \subseteq N$ of $k \equiv |K|$ agents. In section 2.1 we considered the situation in which every $i \in K$ acted independently by choosing x_i so as to maximize u_i. We will refer to this arrangement as decentralization.

Centralization. The equilibrium under decentralization, characterized by (4), will not maximize $u_K \equiv \sum_{i \in K} u_i$ if $\varepsilon \neq 0$. If we take the other agents' actions as given, the coalition K would, in total, benefit the most by selecting $\mathbf{x}_K \equiv \{x_i\}_{i \in K}$ so as to maximize u_K. The Coase theorem is indeed suggesting that the coalition members in K will succeed in maximizing u_K when utilities, as here, are transferable (i.e., every u_i is linear in t_i). Such an arrangement with side transfers might be possible in application (a), when the different agents are producers of fossil fuels, if the agents succeed in establishing a negotiation platform in advance. The oil exporting countries in OPEC, for instance, have a long history with negotiations and collabora-

[8] In equilibrium, however, P does not want to increase x_i above x_i^* because t_i must then be larger, as well. The intuition is that the payment from i to P is, as mentioned, $T_i(x_i) = (c - \underline{c}) x_i - t_i$, implying that i's net profit is $px_i - cx_i$ (with c, and not $\underline{c}$) but i does not like to produce that much when the cost of production is c instead of $\underline{c}$. So, if P increases x_i above x_i^*, the component t_i in $T_i(x_i) = (c - \underline{c}) x_i - t_i$ must increase.

tion. In application (b), with districts, K will maximize u_K if decisions about $\mathbf{x}_K$ are delegated to a central authority, or benevolent manager, that is endowed with (or given) the objective function u_K. In the literature on fiscal federalism, the standard assumption under centralization is indeed that the central government will maximize the sum of the districts' payoffs (Oates, 1972). In the application (c) with retailers, centralization can refer to a merger between the k firms. In this case, it is natural that the merged firm will maximize the sum of profit for the different production units.

In our view, it is consistent with these applications to assume that centralization means that $\mathbf{x}_K$ is set such as to maximize $u_K \equiv \sum_{i \in K} u_i$ at the last stage of the game, at the very same time as every $j \in N \backslash K$ sets x_j. The coalition must, however, make the institutional decision at the beginning of the game, and it cannot, then, commit to anything other than to maximize u_K ex post. Any other commitment would not be renegotiation-proof. The decision to maximize u_K is, obviously, renegotiation-proof.

Thus, centralization implies that the number of decision-making agents in K reduces from k to 1, and overall it reduced from n to $n-(k-1)$. The proofs in the Appendix permit intermediate institutional changes in which the number of decision makers representing K is reduced by any integer Δ: our results hold for any given $\Delta \in \{1, ..., k-1\}$.

Institutional change. It is desirable to let the institutional arrangement be endogenous, but the decision is not made frequently. We find it natural to assume that K evaluates whether to centralize or decentralize at the very beginning of the game. In order to study the incentive to undertake a particular institutional change, in isolation, we will study this decision while we keep fixed the institutional arrangement of the other agents.

If the institutional change increases the equilibrium level of $u_K = \sum_{i \in K} u_i$, we will say that K prefers centralization. Otherwise, we will say that K prefers decentralization. The comparison is not straightforward because if $K \subset N$ centralizes, $i \in N \backslash K$ expects $x_K \equiv \sum_{j \in K} x_j$ to be different from when K decentralizes. Hence, i's choice of x_i will also be influenced by K's institutional choice. This influence must be taken into account by K when K decides whether or not to centralize at the beginning of the game. When side transfers are possible at the first stage of the game, the Coase theorem predicts that K will centralize if and only if the equilibrium u_K is larger under centralization.

When $\varepsilon > 0$, K finds it optimal to reduce x_K if K centralizes. In response, $i \in N \backslash K$ finds it optimal to produce more. When $\varepsilon < 0$, x_K increases when K centralizes, and thus $i \in N \backslash K$ finds it optimal to produce less. In either case, the influence on x_i, $i \in N \backslash K$, is undoing a part of the coalition's effort. If k is large, while $N \backslash K$ is small, the impact on x_i, $i \in N \backslash K$, is less important. If $K = N$, for instance, K always benefits from centralization because that will take all horizontal externalities into account. If k is small, while $N \backslash K$ is large, the negative effects from centralization on x_i, $i \in N \backslash K$, can outweigh the direct beneficial effects when the k merged units internalize the externalities among themselves. Based on this reasoning, it is intuitive that K prefers centralization if and only if k is large relative to n.

Definition. *Coalition K is* large *if and only if $k > \underline{k}(n)$, where*

$$\underline{k}(n) \equiv n + \frac{3}{2} - \sqrt{n + \frac{5}{4}} \in (0, n) . \tag{12}$$

If K is not large, K is small.

Proposition 0. *Consider the situation with no contract. $K \subseteq N$ prefers centralization if and only if K is large.*

Centralization means that the externalities from x_i on the other agents in K are taken into account, but also that agents in $N \backslash K$ will react. Thus, the size of K relative to N will be crucial for the effect of (de)centralization. This intuition is confirmed in the traditional literature on mergers, discussed in the Introduction, and it is also confirmed in our analysis: Proposition 0 follows as a corollary from Proposition 1 below.[9]

3.2 Induced institutional change

We are finally ready to combine the three pieces of our model: Section 2.1, in which we introduced the horizontal competition, Section 2.2, in which we studied the vertical contracts, and Section 3.1, in which we explained how (de)centralization works.

[9] Our results hold regardless of what tie-breaking assumption we impose for the situation in which $k = \underline{k}(n)$.

It is natural to assume that if the principal can contract with the k agents in K when they decentralize, then the principal can also contract with K under centralization. In addition, P offers payment schedules to every $i \in M \backslash K$.

Regarding the timing, for the same reasons as in Section 3.1, we continue to find it natural that K decides whether to centralize at the very beginning. Thereafter, the principal proposes $\mathbf{t}(\mathbf{x})$. (Equivalently, we can permit the principal to revise the contracts after an institutional change.) This timing seems reasonable because while institutional changes take time, vertical contracts can often be adjusted rather rapidly. Given the timing in Figure 1, the equilibrium contracts offered by the principal will be influenced by the institutions that have been formed. When the agents anticipate that the contracts will depend on the institutions, the presence of the principal and the anticipation of the contracts can *induce institutional change*. The induced institutional change can overturn the qualitative predictions stated in Proposition 0.

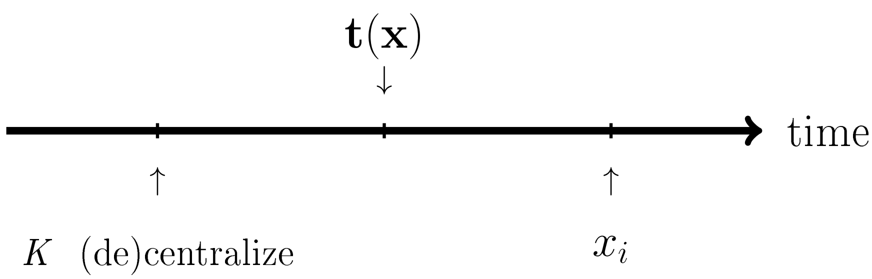


Figure 1: Timing of the game.

The game is solved by backward induction to characterize the unique subgame-perfect equilibrium.

To state the formal results, it will be useful to refer to the following thresholds for the ratio, $r \equiv \varepsilon/\mu$:

$$\widehat{r}_K \equiv 1 - \frac{2(m-k+1)}{n-k+2} - 2\frac{\frac{m}{n+1} - \frac{m-k+1}{n-k+2}}{1 + \frac{n+1}{n-k+2}\sqrt{\frac{1}{k}}}, \tag{13}$$

$$\overline{r}_K \equiv 1 - \frac{2(m-k+1)}{n-k+2} - 2\frac{\frac{m}{n+1} - \frac{m-k+1}{n-k+2}}{1 - \frac{n+1}{n-k+2}\sqrt{\frac{1}{k}}}, \text{ and} \tag{14}$$

$$\widetilde{r} \equiv 1 - \frac{m}{n+1} - \frac{m-k(n)+1}{n-k(n)+2}. \tag{15}$$

It is easy to check that $\widetilde{r} < \widehat{r}_K < \overline{r}_K$ if K is large, but if K is small, the ranking is reversed and $\widetilde{r} > \widehat{r}_K > \overline{r}_K$. Note that $\widetilde{r}$ is independent of k, since it refers to the function (12).

Proposition 1. *There can be contractually induced institutional change.*
(i) Suppose $\varepsilon/\mu > \widetilde{r}$. As is consistent with Proposition 0, a small K always prefers decentralization. For a large K, $\widetilde{r} < \widehat{r}_K < \overline{r}_K$ and K prefers decentralization if and only if $\varepsilon/\mu \in [\widehat{r}_K, \overline{r}_K]$.
(ii) Suppose $\varepsilon/\mu < \widetilde{r}$. As is consistent with Proposition 0, a large K always prefers centralization. For a small K, $\overline{r}_K < \widehat{r}_K < \widetilde{r}$ and K prefers centralization if and only if $\varepsilon/\mu \in [\overline{r}_K, \widehat{r}_K]$.

The threshold $\widehat{r}_K$ is illustrated as the red, upward-sloping line in Figure 2, $\overline{r}_K$ is drawn as the green curves, while $\widetilde{r}$ is the dotted horizontal line. Here, the vertical axis measures ε/μ. Figure 3 draws the same functions, but with a vertical axis that measures $\mu/\varepsilon = 1/r$.

The latter figure is perhaps most intuitive, especially when we start by considering the benchmark case in which $\mu = 0$ (i.e., the horizontal axis). When $k = \underline{k}(n)$, then, K is indifferent between decentralization and centralization. A slightly larger k implies that we move to the right in the figure, i.e., to a region in which K strictly prefers centralization. A slightly smaller k moves us, instead, to the region in which K strictly prefers decentralization. The lessons confirm the insight from Proposition 0.

Instead of moving horizontally, consider a vertical move. Suppose we start from $k = \underline{k}(n)$ and $\mu/\varepsilon = 0$, and increase μ/ε, as in application (a) on fossil fuel production. Then, K strictly prefers decentralization. The intuition is as follows. When $\mu/\varepsilon > 0$, the sign of the horizontal externality is the same as the sign of the vertical externality. P's offer is then inducing $i \in M$ to take an action that is more beneficial also to $j \in K \backslash i$. In other words, the vertical contract between P and $i \in K$ benefits also $j \in K \backslash i$. The positive externality implies that if K had centralized, then K would be willing to go along with the adjustments proposed by P at a lower transfer from P to K. If K stays decentralized, in contrast, then $j \in K \backslash i$ benefits from P's contract with i even if j decides to ignore the payments that P offers j. In other words, the contract between P and i improves j's outside option, and P cannot extract the entire benefit from the contracts for j. Therefore, the transfers and payoffs to the agents in K will be larger if they stay decentralized than if they centralize. The effect on the transfers is then motivating K to decentralize when $\mu/\varepsilon > 0$.

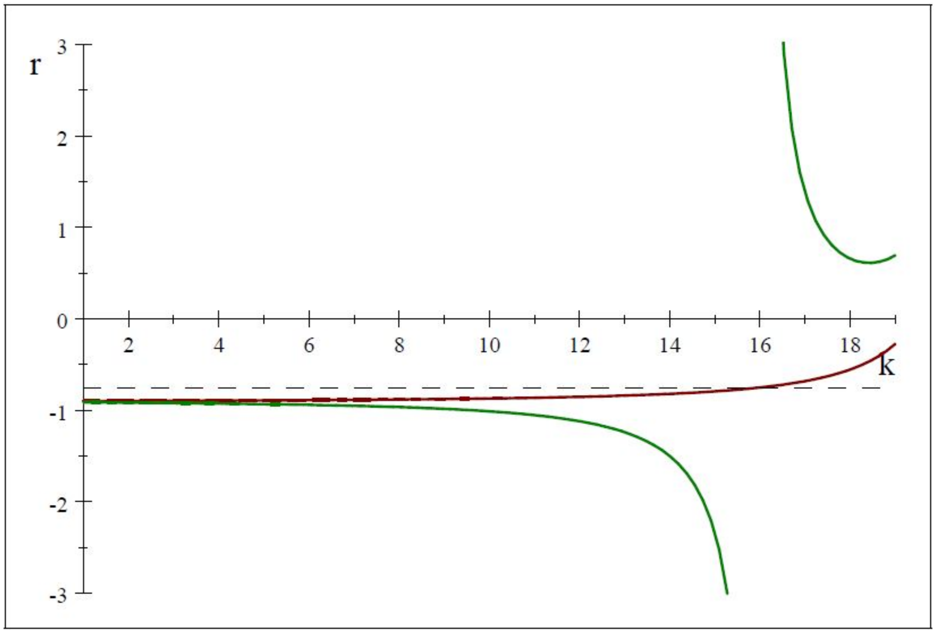


Figure 2: *K prefers decentralization to the left of the curves.*

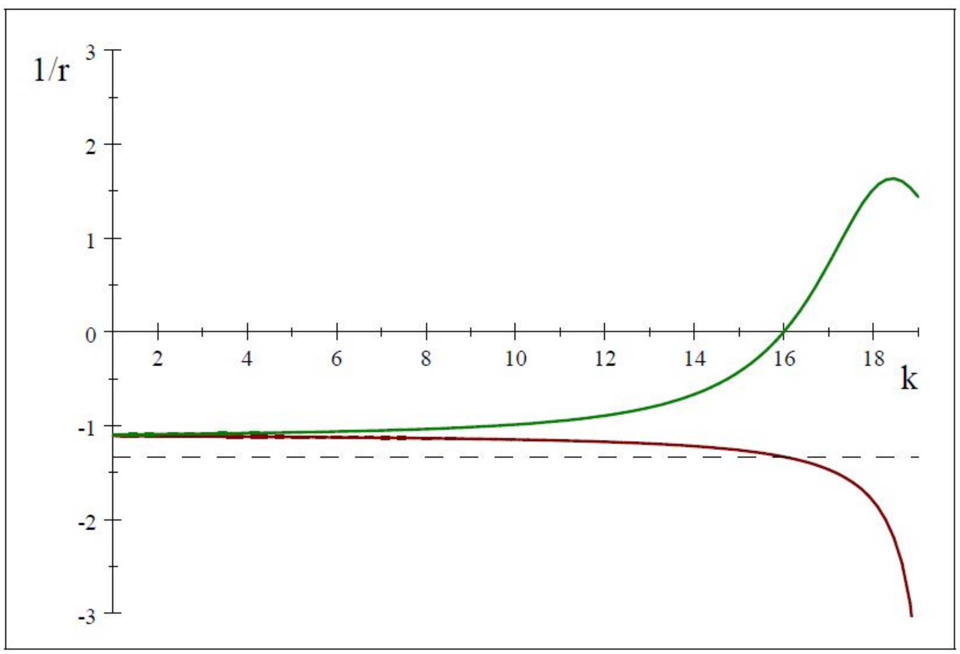


Figure 3: *The thresholds are as in Fig. 2, but now with $\frac{1}{r}$ on the vertical axis. In both figures, curves are drawn for $m = n = 19$, so that $k(n) = 16$ is an integer.*

When $k \approx \underline{k}(n)$, remember, K would be indifferent between the two arrangements in the absence of any contract. If k is larger than $\underline{k}(n)$, then there is a trade-off: The horizontal externality motivates the large K to centralize, but K is also tempted to decentralize to improve the vertical contracts. So, if $k > \underline{k}(n)$, it is not sufficient that $\mu/\varepsilon > 0$ for K to prefer decentralization: μ/ε might need to be quite large for this to be preferred by K.

Suppose now that we start from $k = \underline{k}(n)$ and $\mu/\varepsilon = 0$, and decrease μ/ε, as in applications (b) or (c). Then, K strictly prefers centralization. When $\mu/\varepsilon < 0$, the horizontal externality has the opposite sign of the vertical externality. Thus, the vertical contract between P and $i \in K$ is harmful for $j \in K \backslash i$. If K stays decentralized, then the contract between P and $i \in K$ will be harmful for $j \in K \backslash i$ also if j ignores the payment schedule offered by K. P will not need to compensate j for this harm. If the two merge, instead, then the central authority in K will not go along with the proposal from P, unless P offers larger transfers. Therefore, if $k \approx \underline{k}(n)$ and $\mu/\varepsilon < 0$, K prefers centralization in order to improve the offer from P. If $k < \underline{k}(n)$, K faces a trade-off: centralization improves the offer from P, but decentralization induces the other agents, $i \in N \backslash K$, to adjust their quantities in the direction preferred by K. (This is why a small K prefers decentralization when $\mu = 0$.) Thus, if k is much smaller than $\underline{k}(n)$, it is necessary but insufficient that $\mu/\varepsilon < 0$ for K to prefer centralization: μ/ε might need to be quite a lot smaller than 0 for this institution to be preferred by K.

The lesson is that for applications in which the horizontal and the vertical externalities have the same sign, as in the pollution example (a), then P's concern might be that the anticipation of contracts can motivate large coalitions to decentralize in order to pressure P to offer larger payments. For applications in which the horizontal and the vertical externalities have opposite signs, as in (b) deforestation (where $\varepsilon < 0 < \mu$) or (c) retailers (where $\mu < 0 < \varepsilon$), then induced institutional change will instead take the form of small coalitions that centralize in order to improve the contractual terms offered by P.

There is a limit to these lessons, however, because P's contract is, in itself, changing the equilibrium $\partial u_i / \partial(-x_j)$. If $\mu/\varepsilon < 0$ is much below 0, then P asks for such large changes in the x_j's that $\partial u_i / \partial(-x_j)$ changes sign. For example, if $\varepsilon < 0$, as in application (b), then a large $\mu > 0$ implies that P pays for a large reduction in the x_j's. When the x_j's become very small, and M is large, then i's best response, if i ignores the payments from P, is a large x_i^0. When x_i^0 increases, (6) shows that $\partial u_i / \partial(-x_j)$ can change from being a negative number (when $s_i < 0$) to being a positive number. With this change in the sign of $\partial u_i / \partial(-x_j)$, μ and $\partial u_i / \partial(-x_j)$ obtain the same sign, just as if $\mu/\varepsilon > 0$ were large. When $\mu/\varepsilon > 0$ is large, Figure 3 shows that K

prefers decentralization. This logic explains (rather technically) why K prefers decentralization also if $\mu/\varepsilon < 1/\widehat{r}_K$ in Figure 3, implying $\varepsilon/\mu \in [\widehat{r}_K, 0]$ in Figure 2. In the application (b) on deforestation, the intuition is that when P pays the districts to conserve a lot, the timber price becomes so high that each district i becomes tempted to sell the timber, rather than to spend resources on monitoring the forest. With that temptation, i naturally benefits when x_j is further reduced and the timber price increased. When a district, then, benefits from P's offer to another district, K might prefer decentralization to obtain a better offer.[10]

3.3 Effects on the principal's payoff

If ε/μ is large, it is natural that P benefits if a coalition $K \subseteq M$ centralizes, because the central authority will internalize the positive externalities associated with the contract. With this internalization, P can lower the payments without violating the incentive constraints. If ε/μ is negative and small, it is intuitive that P is harmed if a coalition $K \subseteq M$ centralizes, because when the central authority internalizes negative contractual externalities, P must offer more.

The induced institutional changes predicted by Proposition 1 go against the principal preference. The anticipation of contracts can induce K to decentralize when ε/μ is large, or centralize when ε/μ is small, as is proven in the Appendix.

Proposition 2. *Any induced institutional change, predicted by Proposition 1, makes the principal worse off.*

The Appendix provides details. The result is quite natural, given the intuition that the agents benefit from an induced institutional change – that overturns what the agents would have preferred in the absence of the contracts – if and only if such a change improves P's offer. An improved offer here refers to larger transfers relative to the sacrifices expected from the agents. Naturally, larger transfers, relative to the sacrifices, harm the principal.

[10] This logic, which also explains why $\widehat{\theta}_K < 0$ in Figure 2, requires M to be large. If, instead, M is small, then, when P's contract reduces x_M, the associated increase in $x_{N \setminus M}$ is so important that the best response for $i \in M$ (if i ignores P's payments) is a smaller x_i^0. With a smaller x_i^0, (6) shows that $\partial u_i / \partial (-x_j)$ can change from being a positive number to being a negative number. Thus, if $e/\mu > 0$ is close to zero, P's contract can imply $\partial u_i / \partial (-x_j) < 0$. This logic is consistent with the fact that $\widehat{\theta}_K > 0$ if M is small, and then a small K might prefer centralization even if $\mu/e \geq 0$, as long as $e/\mu \in \left[0, \widehat{\theta}_K\right]$.

3.4 Gross effects on quantities and harm

Induced institutional change will also influence the equilibrium quantities, the x_i^*'s, and thus the harm, h. There are two reasons for this influence.

First, remember that a large K preferred centralization, in Section 2.1, in order to move x_N^0 in K's preferred direction (i.e., to reduce x_N^0 when $\varepsilon > 0$, and to increase x_N^0 when $\varepsilon < 0$). If a large K is, instead, induced to decentralize, then the quantities will move in the opposite direction. This outcome is harmful for P if μ has the same sign as ε.

Second, the induced institutional change predicted by Proposition 1 is motivated by the improved offers that P must make to satisfy the incentive constraints. Because improved offers mean larger payments in return for smaller adjustments, any induced institutional change makes it more expensive for P to encourage her preferred adjustment in quantities. The larger expense means, in turn, that P buys less in terms of adjustments. As a consequence, the harm h, experienced by P, increases.

Proposition 3. *Consider $K = M$. Any induced institutional change, predicted by Proposition 1, increases h. Moreover, an induced institutional change implies that x_N^* and x_M^* increase but x_L^* decreases if $\mu > 0$, while, if $\mu < 0$, x_N^* and x_M^* decrease but x_L^* increases.*

Thus, P's presence is likely to motivate institutional change that leads to the reverse of what P attempts to achieve. In other words, there is a contractual rebound effect.[11]

3.5 Net effects on quantities and harm

The above findings lead to the question: Can the rebound effect of induced institutional change make the contracts counterproductive? That is, can the anticipation of the offers from P induce an institutional change that in itself causes more harm than the reduction in harm motivated by P's offers? If so, P's presence is clearly doing more harm than good and P would have preferred to commit to abstaining from offering contracts, if such a commitment were feasible.

Proposition 4. *Contracts can be counterproductive and increase h:*

[11] The Appendix shows that $K = M$ is not necessary for this result to hold, but the result is somewhat more complicated to describe without this restriction.

(i) Suppose $\varepsilon/\mu > \widetilde{r}$. When there is an induced institutional change, i.e., a large K decentralizes, then $h > h^0$ if:

$$\varepsilon/\mu > \overrightarrow{r}_M \equiv 2\frac{m(2+n-k)}{(k-1)(1+n)} > 0.$$

(ii) Suppose $\varepsilon/\mu < \widetilde{r}$. When there is an induced institutional change, i.e., a small K centralizes, then $h > h^0$ if:

$$\varepsilon/\mu < \overleftarrow{r}_M \equiv -2\frac{(1+n)(m-k+1)}{(k-1)(2+n-k)} < 0.$$

In both cases, $x_N^* > x_N^0$, $x_M^* > x_M^0$, and $x_L^* < x_L^0$, if $\mu > 0$. The reverse holds when $\mu < 0$.

The results are intuitive in light of the insight developed above. When K is large and the vertical and the horizontal externalities have the same signs, the contracts are counterproductive if they motivate K to decentralize, even though they do not dramatically influence the quantities (which is the case when ε/μ is large).

The analogous argument holds when K is small and the externality is negative.

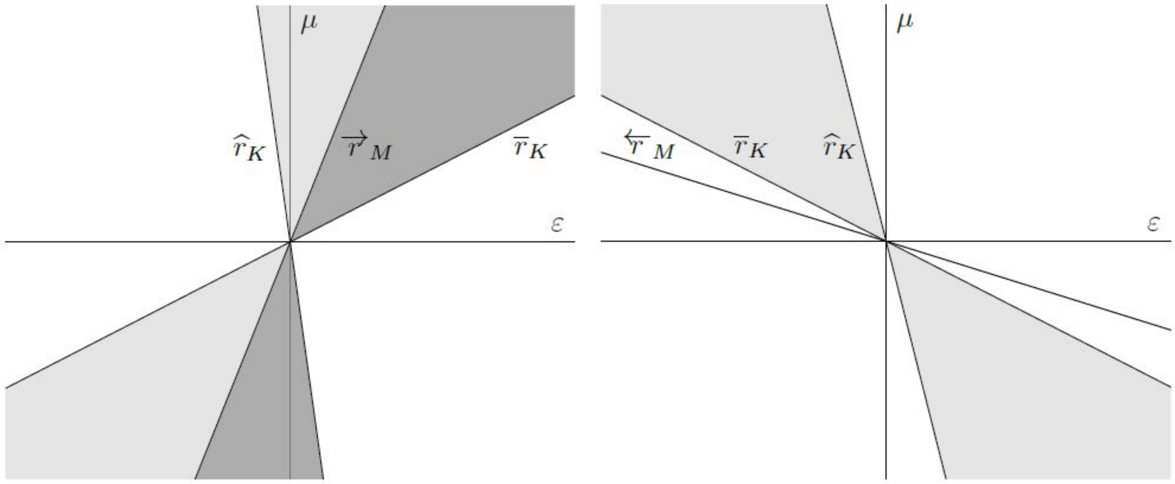
4 Examples and applications

4.1 Quantitative examples

The reader can easily develop numerical assessments by using the formulae above. As a basis for the discussion in the next subsection, we here present the simplest assessment that permits small as well as large coalitions: the case in which $m = n = 3$.

Example A: Large K. To study the incentives for a large group of agents to centralize, consider K of size $k = 3$. This case is interesting because Proposition 0 states that this K would prefer to centralize in the absence of contracts. With centralization, the contractual outcome would be "first best" since P would capture and internalize all the surplus. The question is therefore whether (or when) K rather prefers decentralization because contracts are anticipated. By applying Propositions 1–4, we find:

$$\begin{aligned} \text{If } k &= m = n = 3, \text{ then:} \\ \underline{r}_K &= -0.25 < \widehat{r}_K = 3/2 - \sqrt{3} \approx -0.23 < \overrightarrow{r}_M = 1.5 < \overline{r}_K = 3/2 + \sqrt{3} \approx 3.23. \end{aligned}$$



(a) K Large (Example A). (b) K Small (Example B).

Figure 4: Examples.

The left panel in Figure 4 illustrates in (light and dark) grey the parameter region in which a large K prefers to decentralize because of the contracts. In the dark grey region, the effect of the induced institutional change on x_M and h outweighs the effect of the contracts themselves, and thus the harm is larger than if P had been absent from the game. This situation is evidently not only a theoretical possibility, but also arises for a large set of parameters. Note that the situation is symmetric around the origin, so that every conclusion that holds for some ε and μ also holds for $-\varepsilon$ and $-\mu$.

Figure 4: *Left panel is for K large (Example A). Right panel is for K small (Example B).*

Example B: Small K. To study the incentives for a small group of agents to centralize completely, consider K of size $k = 2$. When $n = 3$, $k = 2$ is actually a "small" group, according to our definition, because $\underline{k}(3) = 2.4$. This case is interesting because Proposition 0 states that K would prefer decentralization in the absence of contracts. The question is therefore whether (or when) the agents prefer centralization when P is present in the game. By applying Propositions 1–4, we find:

$$\begin{aligned} \text{If } k &= 2 < m = n = 3, \text{ then:} \\ \underline{r}_K &= 0.17 > \widehat{r}_K = \sqrt{2} - 11/6 \approx -0.419 > \\ \overline{r}_K &= -\sqrt{2} - 11/6 \approx -3.25 > \overleftarrow{r}_M = -16/3 \approx -5.33. \end{aligned}$$

The right panel in Figure 4 illustrates in grey the parameter set under which a small K prefers to centralize because of the contracts. In this example, the effect of induced institutional change would outweigh the effect of the contracts themselves only if $|\varepsilon/\mu|$ were so large that the induced institutional change would actually not occur. Once again, the situation is symmetric around the origin, so that every conclusion that holds for some ε and μ also holds for $-\varepsilon$ and $-\mu$.

The examples confirm the finding that when ε/μ is large, P's concern should be that a large K might decentralize. When ε/μ is small, P's concern should be that a small K might centralize.

4.2 Implications for the applications

Distinguishing between the incentives of small and large coalitions can be helpful to categorize our results. However, this categorization is hardly helpful in applications, because the principal will not take as given k, but instead she will take the externalities as given. Therefore, a more useful categorization may be in terms of ε and μ, and Propositions 1–4 thus refer to these parameters. We can shed further light on the applications of our model by connecting them more carefully to the above results. Although a serious empirical test would have to await future research, this section is briefly returning to the four applications presented in Section 2.3.

(a) Pollution. When a regulator attempts to reduce pollution, a typical concern is "leakage," that is, that other producers will pollute more, given the general equilibrium effects. In addition, we show that an additional problem can be induced institutional change.

These problems uncover the limits of buying coal with the intention of keeping fossil fuels in the ground, as analyzed by Harstad (2012). For this application, we have $\varepsilon > 0$ and $\mu > 0$. As in application (a) and Example A, the concern is that a large group of fossil-fuel producers may have an interest in ending their coordination and cooperation (i.e., in decentralizing authority), just to obtain larger transfers from P. Thus, if the climate coalition were seeking to compensate major oil exporters (who are members of OPEC) to motivate less production, the concern could be that OPEC could be dismantled and abandoned. If that were to happen, the climate coalition would find it more expensive to reduce pollution.

If the regulator instead subsidizes producers of natural gas, competing with foreign coal producers, then it is realistic that $\mu < 0$. In this case, the regulator would like to motivate increased production that can replace coal. It has been argued that the regulator of Norwegian natural gas production should (or does) consider this possibility. The concern in this situation should be that a small group of gas producers might collude, as in Example B. Such collusion would reduce the quantity the members would produce, and the regulator would need to increase the subsidies to obtain production expansions. In the Norwegian example, this concern seems justified since the natural gas industry is indeed dominated by a few large companies.

In both these cases – i.e., whether P regulates coal or subsidizes gas producers – the anticipation of contracts can lead to institutional changes that reduce the effectiveness of the policy so that total emissions increase. The effects of such institutional changes come in addition to the traditional carbon leakage effects. The regulator can benefit from taking the effects into account.

(b) Deforestation. Our second application is motivated by the conservation contracts better discussed in Harstad and Mideksa (2017). When tropical deforestation is illegal, conservation in one jurisdiction can be harmful for other jurisdictions because of increased pressure from illegal logging. When the externality from j's conservation to i's payoff is negative, while $\mu > 0$, so that P pays to conserve, P's main concern should be that a small coalition of jurisdictions will find it optimal to centralize authority with the intention of obtaining larger transfers, as in Example B.

This result is consistent with the empirical evidence. Even in situations and countries in which decentralization reduces deforestation, Phelps et al. (2010) and Larson and Soto (2008) argue that REDD+ contracts can motivate recipient countries to centralize forest management. Similarly, Bayrak and Marafa (2016:5) fear that "Governments could be inclined to recentralize their forest management systems ... however, bureaucracy and corruption could render a national approach inefficient and counter effective." In line with our analysis, Ribot et al. (2006) show that such recentralization has indeed occurred in Senegal, Uganda, Nepal, Indonesia, Bolivia, and Nicaragua. It is clearly important to understand the motivation for these institutional changes to develop more effective REDD+ agreements in the future.

Despite the differences from the first application, the substantial take-away lesson remains robust: Induced institutional change will harm the principal, reduce the contracts' effectiveness, and possibly lead to more harm than if there were no contracts. If that situation arises, the contracts may do more harm than good.[12]

Given these results, it is worthwhile to consider the Norwegian government's rejection of cooperating with the region Madre de Dios in southeastern Peru (this case is discussed in Harstad and Mideksa, 2017). This rejection seems rational if one believes that the primary motivation for deforestation is that jurisdictions seek to profit (i.e., $\varepsilon > 0$). The concern is then, as in Example A, that decentralization (to regions, for example) will occur as a way of eliciting larger compensations, even though less will be conserved as a result. However, given that a large part of deforestation in Peru is illegal,[13] it might be reasonable to assume that conservation in one region leads to more pressure from illegal loggers in the others (i.e., $\varepsilon < 0$). In this case, the donor's primary concern should be that jurisdictions collude rather than decentralize. It follows that an agreement with Madre de Dios should be embraced when the main problem is illegal logging, even though such an agreement will be rejected if the donor presumes logging to be driven by the jurisdictions' search for profit.

(c) Retailers. A quite different situation arises when an upstream supplier (or parent company) can provide useful inputs for downstream firms. While we have $\varepsilon > 0$ for the firms, we have $\mu < 0$ because the supplier's cost-saving input makes it beneficial for the supplier to increase the retailers' sales, as discussed in Section 2.3. The competitors, in response, will reduce their quantities when the supplier's retailers become more competitive. In this situation, the concern is not that large groups of firms will end cooperation or coordination (i.e., decentralize), but that small groups of firms will collude (as in Example B) as a way to obtain more attractive offers from the supplier.

This prediction is in line with findings in the literature discussed in Section 1: The empirically observed consolidation in the U.K. grocery retail industry and the mergers among hospitals in Massachusetts are in line with the model because the horizontal and the vertical externalities

[12] In principle, the REDD+ application can also motivate the possibility that both μ and ε would be negative. This situation arises when the donor recognizes that more logging in the contracted-with districts replaces illegal logging of forests that have larger conservation values. However, this case is very much like the case in which both ε and μ are positive, and the primary concern for P would be that a large group of districts will decentralize or end coordination and cooperation in order to obtain more compensation, just as in Example A.

[13] Goncalves et al. (2012) report that 80% of the total Peruvian timber exports is from illegal logging.

have opposite signs in the supplier–retailers application. If the externalities had the same sign instead, the predicted induced institutional change would have been decentralization.

The predicted institutional change will always harm the supplier and it is likely to reduce the quantity produced. This reduction can potentially outweigh the direct effect of the vertical contracts. In contrast to the traditional collusion emphasized by economists and antitrust agencies, the contractually induced institutional changes are made by retailers not to obtain a higher price from consumers, but to obtain better offers from their common supplier. Nevertheless, the induced institutional changes harm consumers and thus, the changes should be of interest to antitrust agencies.

5 Concluding remarks and future research

This article combines a simple model of collusion with a model of vertical contracts. The contracts must be more attractive under centralization if and only if the externality facing the principal and that facing other agents have opposite signs. In this situation, agents may prefer to centralize to obtain better deals. If the two externalities have the same signs, the agents are more likely to decentralize to obtain the best contract. The induced institutional changes always harm the principal, and the negative effects can outweigh the direct effects of the contracts themselves.

Our findings have important implications for applications such as the regulation of pollution, payments for forest conservation, and vertical contracts when there are multiple retailers. The results suggest that environmental policies may face problems in addition to those posed by emission leakages. There is some evidence suggesting that institutions are endogenous in countries that are offered agreements on REDD+, but deforestation and conservation are research topics that are yet to be satisfactorily investigated empirically as well as theoretically. Vertical contracts in supplier–retailers relationships are less effective, we predict, with contractually induced mergers. All these predictions are empirically testable and should be matched with data in future research to validate the importance of the results.

We have followed the early literature on industrial organization by assuming a linear demand function. This approach resulted in a linear-quadratic utility function for every agent, and simplified the expressions of the solutions. The subsequent literature on industrial organization has relaxed the linear-demand assumption, but the essence of the results regarding mergers has been shown to hold.[14]

The Cournot game is also implying that actions are strategic substitutes. We believe, however, that our main results generalize to settings in which the agents' actions are strategic complements, because the sign of the externality does not depend on substitutability. Consider, for instance, Bertrand price competition. If one producer sets a higher price, other producers will also charge higher prices (thus, actions are strategic complements). If a principal pays one producer to reduce (raise) the price, the vertical and the horizontal externalities have opposite (same) signs, and the principal can pay less unless the agents (de)centralize. To improve the offer, agents should (de)centralize when the externalities have opposite (same) signs, exactly as our theory predicts.

Nevertheless, our model is simple and stylized and should be generalized in future research. An important limitation is that we have considered the incentive to centralize among only a single set of members. This assumption ought to be relaxed in two ways. If we permit the reforming coalition to include agents which the principal cannot contract with, then this coalition might realize that if it centralizes, the principal becomes less concerned with leakage and she might offer more ambitious contracts. This is beneficial for the coalition if and only if the two externalities have the same sign. In this case, the principal can benefit when her presence motivates the non-members to centralize. If the externalities are of opposite signs, the coalition might decentralize to make the contracts less ambitious. This induced institutional change can harm the principal.

After one institutional change, other coalitions may also reform. Future research should consider the setting in which multiple subsets of agents can agree on institutional changes. In reality, jurisdictions are divided among different countries, and each country may decide on whether to centralize or decentralize. In this setting, the institutional changes can be strategic

[14] One illustration of this robustness is the analysis by Gaudet and Salant (1991), mentioned above, which considers when a subset of firms would benefit from marginally reducing their production levels. Gaudet and Salant (1991:658) find that "a marginal contraction is strictly beneficial (strictly harmful) if and only if the number of firms in the designated subset exceeds the 'adjusted' number of firms outside it by strictly more (strictly less) than one. The adjustment factor is unity when cost and demand functions are linear but, more generally, depends on the convexity of the cost and demand curves."

complements, because when one subset centralizes, the total number of decision makers declines, and this decline makes it more likely that another subset will also prefer to centralize. This complementarity can lead to multiple equilibria. Future research should search for the best equilibrium and for how the principal can induce the agents to coordinate on such an equilibrium.

Appendix: Proofs

Remark on generalizations: The proofs permit the possibility that K reduces the number of producers from k to $k-\Delta$. By permitting $\Delta \in \{1, ..., k-1\}$, we allow for intermediate cases in-between full centralization and full decentralization. We here take both k and Δ as given numbers, and investigate when the sum of payoffs for the agents in K increases when the number of decision makers in K is reduced from k to $k-\Delta$. The propositions in the text follow when we limit attention to $\Delta = k-1$.

Definition. *Coalition K is* large *if and only if:*

$$\underline{r}_K \equiv 1 - \frac{k}{n+1} - \frac{k-\Delta}{n-\Delta+1} < 0. \tag{16}$$

If K is not large, K is small.

With $\Delta = k-1$, the inequality can be written as $k \geq \underline{k}(n)$, as in the definition in Section 3.1.

Proof of Lemma 1. *Best responses:* From the first-order condition (f.o.c.) of (2) with respect to (w.r.t) x_i, we obtain:

$$x_i = \frac{w\bar{p} - wax_N - c - s_i}{wa} = \frac{w\bar{p} - wax_M - wax_L - c - s_i}{wa}.$$

The market equilibrium is, therefore:

$$\begin{aligned}
x_M &= \frac{mw\bar{p} - mwax_M - mwax_L - mc - s_M}{wa}, \\
x_L &= \frac{lw\bar{p} - lwax_M - lwax_L - lc - s_L}{wa}, \\
x_N &= \frac{nw\bar{p} - nwax_N - nc - s_M - s_L}{wa} = \frac{nw\bar{p} - nc - s_M - s_L}{wa(1+n)}, \\
x_i &= \frac{w\bar{p} - c + s_M + s_L - s_i(1+n)}{wa(1+n)}.
\end{aligned}$$

The externality is:

$$\frac{\partial u_i\left(x_i, x_{N\setminus i}\right)}{\partial(-x_j)} = awx_i + s_i = \frac{w\bar{p} - c + s_N}{1+n}.$$

Non-participants' response: From the best-response function, when we add up all the x_i's, $i \in L$, we obtain:

$$x_L(x_M) = \frac{lw\bar{p} - lwax_M - lc - s_L}{wa(1+l)}, \tag{17}$$

where we have abused notation by writing the function as $x_L(x_M)$.

Payoffs: A participant's payoff as a function of x_i and the contracts is:

$$
\begin{aligned}
u_i\left(\mathbf{x}^*_{M\setminus i};x_i\right) &= w\left(\bar{p}-ax_{N\setminus i}-ax_i\right)x_i - cx_i - s_i x_L\left(x_M\right) - s_i x_{M\setminus i} - s_i x_i \\
&= \left[\frac{w\bar{p}-wax_{M\setminus i}+lwax^*_i+lc+s_L}{1+l} - c - s_i - wax_i\right]x_i - s_i x_L\left(x_M\right) - s_i x_{M\setminus i}.
\end{aligned}
$$

The f.o.c. of this objective w.r.t. x_i gives us i's best response, if i decides to respond optimally to the quantities chosen by the others, given the contracts:

$$
x_i\left(\mathbf{x}^*_M\right) = \frac{w\bar{p}-wax_{M\setminus i}+lwax^*_i+lc+s_L}{2aw\left(1+l\right)} - \frac{c+s_i}{2aw}.
$$

The corresponding utility is i's "outside option":

$$
\begin{aligned}
u_i\left(\mathbf{x}^*_{M\setminus i};x_i\left(\mathbf{x}^*_M\right)\right) &= wax_i^2 - s_i x_L\left(x_M\right) - s_i x_{M\setminus i} \qquad (18)\\
&= \frac{1}{4aw}\left(\frac{w\bar{p}-wax_{M\setminus i}+lwax^*_i+lc+s_L}{1+l} - \left(c+s_i\right)\right)^2 - s_i x_L\left(x_M\right) - s_i x_{M\setminus i}.
\end{aligned}
$$

The transfer: t_i must equal i's reduction in payoff from accepting the contract x^*_i, relative to i's outside option. Thus:

$$
\begin{aligned}
t_i\left(\mathbf{x}\right) &= u_i\left(\mathbf{x}^*_{M\setminus i};x_i\left(\mathbf{x}^*_M\right)\right) - u_i\left(\mathbf{x}^*_M\right) \\
&= \left(\frac{w\bar{p}-wax_{M\setminus i}+lwax^*_i+lc+s_L}{2aw\left(1+l\right)} - \frac{c+s_i}{2aw}\right)2aw\left(x_i-x^*_i\right) - wa\left(x_i^2-\left(x^*_i\right)^2\right) \\
&= 2aw\left(x_i^2-x_i x^*_i\right) - wa\left(x_i^2-\left(x^*_i\right)^2\right) = wa\left(x_i-x^*_i\right)^2 \\
&= wa\left(\frac{w\bar{p}-wax_M+lc+s_L}{2aw\left(1+l\right)} - \frac{\widetilde{x}^*_i}{2}\right)^2\text{, where} \qquad (19)\\
\widetilde{x}^*_i &\equiv x^*_i + \left(c+s_i\right)/aw.
\end{aligned}
$$

P's objective function becomes symmetric in all the $\widetilde{x}^*_i$'s, and the optimal $\widetilde{x}^*_i$'s will be the same for every $i \in M$:

$$
\begin{aligned}
u_P &= -d_M x_M - d_L\frac{lw\bar{p}-lwax_M-lc-s_L}{wa\left(1+l\right)} - wa\sum_{i\in M}\left(\frac{w\bar{p}-wax_M+lc+s_L}{2aw\left(1+l\right)} - \frac{\widetilde{x}^*_i}{2}\right)^2 \\
&= -d_M x_M - d_L\frac{lw\bar{p}-lwax_M-lc-s_L}{wa\left(1+l\right)} - wa\sum_{i\in M}\left(\frac{w\bar{p}+nc+s_N}{2aw\left(1+l\right)} - \frac{1+n}{1+l}\frac{\widetilde{x}^*_i}{2}\right)^2.
\end{aligned}
$$

The f.o.c. w.r.t. $\widetilde{x}^*_i$ is identical for every $i \in M$. Taking into account that $x^*_i \equiv \widetilde{x}^*_i - \left(c+s_i\right)/aw$, we obtain:

$$
\begin{aligned}
0 &= -d_M + \frac{l}{1+l} d_L + aw \frac{1+n}{1+l} \left(\frac{w\bar{p} + nc + s_N}{2aw\,(1+l)} - \frac{1+n}{1+l} \frac{\widetilde{x}_i^*}{2} \right) \Rightarrow \\
\frac{aw}{2} \left(\frac{1+n}{1+l} \right)^2 \widetilde{x}_i^* &= \frac{-\mu}{1+l} + \frac{1+n}{1+l} \left(\frac{w\bar{p} + nc + s_N}{2\,(1+l)} \right) \Rightarrow \\
\widetilde{x}_i^* &= \frac{w\bar{p} + nc + s_N}{aw\,(1+n)} - 2 \frac{1+l}{aw\,(1+n)^2} \mu \Rightarrow \qquad (20) \\
x_i^* &= \frac{w\bar{p} + nc + s_N}{aw\,(1+n)} - 2 \frac{1+l}{aw\,(1+n)^2} \mu - (c + s_i)\,/aw \\
&= \frac{\varepsilon}{aw\,(1+n)} - 2 \frac{1+l}{aw\,(1+n)^2} \mu - \frac{s_i}{aw}.
\end{aligned}
$$

Note that the second-order condition always holds. If we sum over every $i \in M$, we obtain:

$$
\begin{aligned}
x_M^* &= \frac{m\varepsilon}{aw\,(1+n)} - 2m \frac{1+l}{aw\,(1+n)^2} \mu - \frac{s_M}{aw}, \text{ while} \qquad (21) \\
x_L\,(x_M^*) &= \frac{lw\bar{p} - lc - s_L}{wa\,(1+l)} - \frac{l}{1+l} x_M = \frac{lw\bar{p} - lc - s_L}{wa\,(1+l)} - \frac{l}{1+l} \frac{\sum_{i \in M} \varepsilon_i}{aw\,(1+n)} \qquad (22) \\
&\quad + \frac{l}{1+l} 2m \frac{1+l}{aw\,(1+n)^2} \mu + \frac{l}{1+l} \frac{s_M}{aw} \\
&= \frac{l\varepsilon}{aw\,(1+n)} - \frac{s_L}{wa} + \frac{2ml}{aw\,(1+n)^2} \mu. \qquad (23)
\end{aligned}
$$

The transfer is, from (19) and (20):

$$
\begin{aligned}
t_i &= aw \left(\frac{w\bar{p} - wax_M + c_L + s_L}{2aw\,(1+l)} - \frac{\widetilde{x}_i^*}{2} \right)^2 \\
&= aw \left(\frac{w\bar{p} + c_N + s_N}{2aw\,(1+l)} - \frac{1+n}{1+l} \frac{\widetilde{x}_i^*}{2} \right)^2 = \frac{1}{aw} \left(\frac{\mu}{1+n} \right)^2.
\end{aligned}
$$

So, the sum of transfers is simply:

$$
t_M = \frac{m}{aw} \left(\frac{\mu}{1+n} \right)^2.
$$

The total harm is:

$$
\begin{aligned}
h &= d_M x_M + d_L \left(\frac{lw\bar{p} - lc - s_L}{wa\left(1+l\right)} - \frac{l}{1+l} x_M \right) \\
&= x_M \left(d_M - \frac{l}{1+l} d_L \right) + d_L \left(\frac{lw\bar{p} - lc - s_L}{wa\left(1+l\right)} \right) \\
&= \frac{\mu/aw}{1+l} \frac{m\varepsilon}{1+n} - \frac{2m\mu^2}{aw\left(1+n\right)^2} - \frac{\mu/aw}{1+l} s_M + d_L \left(\frac{lw\bar{p} - lc - s_L}{wa\left(1+l\right)} \right).
\end{aligned}
$$

P's payoff is then:

$$
u_P = \frac{m\mu^2}{aw\left(1+n\right)^2} - \frac{\mu/aw}{1+l} \frac{m\varepsilon}{1+n} + \frac{\mu/aw}{1+l} s_M - d_L \left(\frac{lw\bar{p} - lc - s_L}{wa\left(1+l\right)} \right). \quad QED \tag{24}
$$

Lemma 2. *A participant's equilibrium payoff is:*

$$
\frac{1}{aw} \left(\frac{\varepsilon}{n+1} + \frac{m-1-l}{\left(1+n\right)^2} \mu \right)^2 - \frac{\left(w\bar{p} - c\right) s_i}{wa}. \tag{25}
$$

Proof. From the equations above, i's payoff is:

$$
\begin{aligned}
u_i\left(\mathbf{x}^*_{M\setminus i}; x_i\left(\mathbf{x}^*_M\right)\right) &= \frac{1}{4aw} \left(\frac{w\bar{p} - wax_{M\setminus i} + lwax^*_i + lc + s_L}{1+l} - \left(c + s_i\right) \right)^2 - s_i x_L - s_i x_{M\setminus i} \\
&= \frac{1}{4aw} \left(\frac{w\bar{p} - \left(m-1-l\right) wa\widetilde{x}_i + \left(m-1\right) c + s_N - \left(1+l\right) c}{1+l} - 2s_i \right)^2 \\
&\quad - s_i x_L - s_i x_{M\setminus i} \\
&= \frac{1}{4aw} \left(\frac{\varepsilon - \left(m-1-l\right) wa\left(\widetilde{x}_i - c/aw\right)}{1+l} - 2s_i \right)^2 - s_i x_L - s_i x_{M\setminus i} \\
&= \frac{1}{aw} \left(\frac{\varepsilon}{n+1} + \frac{m-1-l}{\left(1+n\right)^2} \mu - s_i \right)^2 - s_i x_L - s_i x_{M\setminus i}.
\end{aligned}
$$

The last two terms, $s_i x_L + s_i x_{M\setminus i}$, can be written as follows:

$$
\left[\frac{ls_i + s_i\left(m-1\right)}{wa\left(1+n\right)} \right] \varepsilon - \frac{2}{aw\left(1+n\right)^2} \mu \left[\left(1+l\right)\left(m-1\right) s_i - s_i lm \right] - \frac{s_i s_L}{aw} - \frac{s_i s_{M\setminus i}}{aw}.
$$

So, combined, i's payoff is $u_i\left(\mathbf{x}^*_{N\backslash i}; x_i\left(\mathbf{x}^*_M\right)\right)$:

$$\begin{aligned}
&= \frac{1}{aw}\left(\frac{\varepsilon}{n+1}+\frac{m-1-l}{(1+n)^2}\mu - s_i\right)^2 \\
&\quad -\left[\frac{ls_i+s_i(m-1)}{wa(1+n)}\right]\varepsilon+\frac{2}{aw(1+n)^2}\mu\left[(1+l)(m-1)s_i - s_i lm\right]+\frac{s_i s_L}{aw}+\frac{s_i s_M\backslash i}{aw} \\
&= \frac{1}{aw}\left(\frac{\varepsilon}{n+1}+\frac{m-1-l}{(1+n)^2}\mu\right)^2-\left[\frac{ls_i+s_i(m+1)}{wa(1+n)}\right]\varepsilon+\frac{s_i s_L}{aw}+\frac{s_i s_{M\backslash i}}{aw}+\frac{(s_i)^2}{aw} \\
&= \frac{1}{aw}\left(\frac{\varepsilon}{n+1}+\frac{m-1-l}{(1+n)^2}\mu\right)^2-\frac{(\varepsilon - s_N)s_i}{wa}. \qquad \mathit{QED}
\end{aligned}$$

Proof of Proposition 0. This result follows from Proposition 1 in the special case where $\mu \to 0$.

Proof of Proposition 1. Analogously to the definition above, we say that M is large if and only if:

$$\underline{r}_M \equiv 1-\frac{m}{n+1}-\frac{m-\Delta}{n-\Delta+1}<0.$$

The payoff of k participating agents is simply k multiplied by (25). If these agents reduce their number by Δ, n and m also decrease by Δ, and the sum of payoffs becomes:

$$\frac{k-\Delta}{aw}\left(\frac{\varepsilon}{n+1-\Delta}+\frac{m-1-l-\Delta}{(1+n-\Delta)^2}\mu\right)^2-\sum_{i\in K}\frac{(w\bar{p}-c)s_i}{wa}.$$

Since $\sum_{i\in K} s_i$ is unchanged in the applications, regardless of whether K (de)centralizes, the merger leads to a larger sum of payoffs than if K does not merge (i.e., k multiplied by (25)), iff:

$$\begin{aligned}
\frac{k-\Delta}{aw}\left(\frac{\varepsilon}{n+1-\Delta}+\frac{m-1-l-\Delta}{(1+n-\Delta)^2}\mu\right)^2 &> \frac{k}{aw}\left(\frac{\varepsilon}{n+1}+\frac{m-1-l}{(1+n)^2}\mu\right)^2 \Rightarrow \\
\frac{n+1}{n+1-\Delta}\sqrt{\frac{k-\Delta}{k}} &> \left|\frac{\varepsilon/\mu+\frac{m-1-l}{1+n}}{\varepsilon/\mu+\frac{m-1-l-\Delta}{1+n-\Delta}}\right| = \left|\frac{\varepsilon/\mu-1+2\frac{m}{1+n}}{\varepsilon/\mu-1+2\frac{m-\Delta}{1+n-\Delta}}\right|. \qquad (26)
\end{aligned}$$

$$\underline{r}_K \equiv 1-\frac{k}{n+1}-\frac{k-\Delta}{n-\Delta+1}$$

Lemma 3: *The left-hand side of (26) is larger than 1 if and only if K is large. In fact:*

$$Q \equiv \frac{n+1}{n+1-\Delta}\sqrt{\frac{k-\Delta}{k}}=\sqrt{1-\frac{\Delta}{k}\left(\frac{n+1}{n-\Delta+1}\right)\underline{r}_K}>1 \Leftrightarrow \underline{r}_K<0.$$

Proof. $\sqrt{\frac{k-\Delta}{k}\left(\frac{n+1}{n-\Delta+1}\right)^2}$ can be written as:

$$\begin{aligned}\sqrt{\frac{1}{k}\frac{(n+1)^2}{n-\Delta+1}\left(\frac{k-\Delta}{1+n-\Delta}\right)} &= \sqrt{1-\left(\frac{1}{k}\frac{(n+1)^2}{n-\Delta+1}-(n+1-k)\frac{1}{k}-1\right)\underline{r}_K} \\ &= \sqrt{1-\frac{\Delta}{k}\left(\frac{n+1}{n-\Delta+1}\right)\underline{r}_K}. \; \|\end{aligned}$$

Substituting for Q in (26), there are three cases to consider:
(a) Suppose $\varepsilon/\mu - 1 + 2\frac{m}{1+n} < 0$. Then, $\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta} < 0$, and (26) can be written as:

$$\begin{aligned}Q &> \frac{\varepsilon/\mu - 1 + 2\frac{m}{1+n}}{\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta}} = 1 + \frac{2\frac{m}{1+n} - 2\frac{m-\Delta}{1+n-\Delta}}{\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta}} \Rightarrow \\ -\left(\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta}\right)(1-Q) &< 2\left(\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right).\end{aligned}$$

This always hold when K is large (since then $Q > 1$, so the l.h.s. is negative). If K is small, this expression becomes:

$$\begin{aligned}\varepsilon/\mu &> \overline{r}_K \equiv 1 - 2\frac{m-\Delta}{1+n-\Delta} - \frac{2}{1-Q}\left(\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right) \\ &= \underline{r}_M - \left(\frac{3-Q}{1-Q}\right)\left(\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right) < \underline{r}_M\end{aligned}$$

when K is small.
(b) Suppose $\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta} > 0$. Then, $\varepsilon/\mu - 1 + 2\frac{m}{1+n} > 0$, and (26) can be written as:

$$\begin{aligned}Q &> \frac{\varepsilon/\mu - 1 + 2\frac{m}{1+n}}{\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta}} = 1 + \frac{2\frac{m}{1+n} - 2\frac{m-\Delta}{1+n-\Delta}}{\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta}} \Rightarrow \\ \left(\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta}\right)(Q-1) &> 2\left(\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right),\end{aligned}$$

which never holds when K is small (since then $Q < 1$, so the l.h.s. is negative), but if K is large, this eq. becomes:

$$\begin{aligned}\varepsilon/\mu &> 1 - 2\frac{m-\Delta}{1+n-\Delta} + \frac{2}{Q-1}\left(\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right) \\ &= \overline{r}_K = \underline{r}_M + \left(\frac{2}{Q-1} + 1\right)\left(\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right) > \underline{r}_M\end{aligned}$$

when K is large (since then $Q > 1$).

(c) Suppose $\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta} < 0 < \varepsilon/\mu - 1 + 2\frac{m}{1+n}$. Then, (26) can be written as:

$$\begin{aligned}
Q &> -\frac{\varepsilon/\mu - 1 + 2\frac{m}{1+n}}{\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta}} = -1 - \frac{2\frac{m}{1+n} - 2\frac{m-\Delta}{1+n-\Delta}}{\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta}} \Rightarrow \\
\left(\varepsilon/\mu - 1 + 2\frac{m-\Delta}{1+n-\Delta}\right) &< -\frac{2\frac{m}{1+n} - 2\frac{m-\Delta}{1+n-\Delta}}{Q+1} \Rightarrow \\
\varepsilon/\mu < \widehat{r}_K &\equiv 1 - 2\frac{m-\Delta}{1+n-\Delta} - \frac{2}{Q+1}\left(\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right) \\
&= \underline{r}_M - \left(\frac{1-Q}{Q+1}\right)\left(\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right),
\end{aligned}$$

which is larger than $\underline{r}_M$ iff K is large.

Finally, note that if $\Delta = k-1$, then $\underline{r}_M$ can be written as:

$$\underline{r}_M \equiv 1 - \frac{m}{n+1} - \frac{m-(k-1)}{n-(k-1)+1},$$

which is increasing in k, and which, by definition, takes the value $\overline{r}$ when $k = k(n)$.

When we summarize the three cases, we arrive at Proposition 1. *QED*

Proof of Proposition 2. P's payoff is (24) and, after the merger, P's payoff is:

$$u_P = \frac{(m-\Delta)\mu^2}{aw(1+n-\Delta)^2} - \frac{\mu/aw}{1+l}\frac{(m-\Delta)\varepsilon}{1+n-\Delta} + \frac{\mu/aw}{1+l}s_M - d_L\left(\frac{lw\bar{p} - c_L - s_L}{wa(1+l)}\right),$$

so the last two terms are unchanged. P's payoff is larger after the merger if and only if:

$$\begin{aligned}
\frac{(m-\Delta)\mu^2}{aw(1+n-\Delta)^2} - \frac{\mu/aw}{1+l}\frac{(m-\Delta)\varepsilon}{1+n-\Delta} &> \frac{m\mu^2}{aw(1+n)^2} - \frac{\mu/aw}{1+l}\frac{m\varepsilon}{1+n} \Rightarrow \\
\frac{\mu\varepsilon}{1+l}\left[\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right] &> \mu^2\left[\frac{m}{(1+n)^2} - \frac{m-\Delta}{(1+n-\Delta)^2}\right] \Rightarrow \\
\frac{\varepsilon}{\mu} &> (1+l)\frac{\frac{m}{(1+n)^2} - \frac{m-\Delta}{(1+n-\Delta)^2}}{\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}} \qquad (27) \\
&= (1+l)\frac{\frac{m(1+n-\Delta)}{1+n} - \frac{(m-\Delta)(1+n-\Delta+\Delta)}{1+n-\Delta}}{m(1+n-\Delta) - (m-\Delta)(1+n)} \\
&= \frac{1+l}{1+n-\Delta} - \frac{m}{1+n} = 1 - \frac{m-\Delta}{1+n-\Delta} - \frac{m}{1+n} = \underline{r}_M.
\end{aligned}$$

However, according to Proposition 2, if K is large, the merger may not take place when $\frac{\varepsilon}{\mu} > \underline{r}_M$ even though K would have merged if P were absent. If K is small, P's presence may lead to a merger but this can happen only if $\frac{\varepsilon}{\mu} < \underline{r}_M$, so exactly when P looses from such a merger. *QED*

Proof of Proposition 3. Proposition 3 follows as a corollary from the more general Lemma

4.

Lemma 4. *If K centralizes, x_N^* and x_M^* decrease but x_L^* increases if and only if*

$$\varepsilon > 2\mu \underline{r}_M,$$

while the reverse holds if $\varepsilon < 2\mu \underline{r}_M$. Consequently:
(i) Suppose K and M are large. Induced institutional change increases h. Moreover, x_N^* and x_M^* increase but x_L^* decreases if $\mu > 0$, while x_N^* and x_M^* decrease but x_L^* increases if $\mu < 0$.
(ii) Suppose K and M are small. Induced institutional change increases h. Moreover, x_N^* and x_M^* increase but x_L^* decreases if $\mu > 0$; while x_N^* and x_M^* decrease but x_L^* increases if $\mu < 0$.

Proof of Lemma 4. First, note that x_L is affected by the merger only through x_M, and that we have $\partial x_L / \partial x_M = -l / (1+l)$.

From (21), we have the expression for x_M. With centralization, it becomes:

$$\frac{(m-\Delta)\,\varepsilon}{aw\,(1+n-\Delta)} - 2\,(m-\Delta)\,\frac{1+l}{aw\,(1+n-\Delta)^2}\mu - \frac{s_M}{aw}.$$

Thus, the merger increases x_M if and only if:

$$\begin{aligned}
\frac{m\varepsilon}{aw\,(1+n)} - 2m\frac{1+l}{aw\,(1+n)^2}\mu &< \frac{(m-\Delta)\,\varepsilon}{aw\,(1+n-\Delta)} - 2\,(m-\Delta)\,\frac{1+l}{aw\,(1+n-\Delta)^2}\mu \Rightarrow \\
\varepsilon\left[\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right] &< 2\mu\,(1+l)\left[\frac{m}{(1+n)^2} - \frac{m-\Delta}{(1+n-\Delta)^2}\right] \Rightarrow \\
\varepsilon &< 2\mu\left[(1+l)\,\frac{\frac{m}{(1+n)^2} - \frac{m-\Delta}{(1+n-\Delta)^2}}{\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}}\right] = 2\mu \underline{r}_M, \qquad (28)
\end{aligned}$$

where the equality in (28) is proven in the proof of Proposition 2. (The expression in the brackets is the same as we have on the right-hand side of (27).)
(i) Suppose K is large. If there is an induced institutional change, it must be decentralization because of P's presence. Decentralization increases x_M when (28) fails, implying $\varepsilon/\mu > 2\underline{r}_M$ when $\mu > 0$, and this always hold under induced institutional changes, according to Proposition 1(i). When $\mu < 0$, (28) fails when $\varepsilon/\mu < 2\underline{r}_M$, and this never holds under induced institutional changes, according to Proposition 1(i). Hence, when $\mu < 0$, an induced institutional change reduces x_M. We know already that x_L moves in the opposite direction from what x_M does, since $\partial x_L / \partial x_M = -l / (1+l) < 0$. The total harm increases when x_M increases if and only if $\mu > 0$, because we can write:

$$\frac{\partial h}{\partial x_M} = d_M - \frac{l}{1+l} d_L = \frac{\mu}{1+l}.$$

(ii) Suppose K and M are small. If there is an induced institutional change, it means that there is a merger because of P's presence. In this case, the institutional change increases x_M when (28) holds, implying $\varepsilon/\mu < 2\underline{r}_M$ when $\mu > 0$, and this always holds under induced institutional changes, according to Proposition 1(ii), because when M is small, then $\underline{r}_M > 0$ and, in this case, $\widehat{r}_K < \underline{r}_M$ implies $\widehat{r}_K < 2\underline{r}_M$. When $\mu < 0$, (28) holds when $\varepsilon/\mu > 2\underline{r}_M$, and this never holds under an induced institutional change when K and M are small, because Proposition 1(ii) states that under an induced institutional change, $\varepsilon/\mu < \widehat{r}_K < \underline{r}_M$, which in turn is smaller than $2\underline{r}_M$ when M is small. The rest of the proof is identical to that of part (i). *QED*

Proof of Proposition 4. We have from (22), as before, that x_L moves in the opposite direction from what x_M does. So, we can start by concentrating on the effects of x_M.
(i) Suppose K is large. If P is not present, K prefers to centralize, and x_M follows from (21) if m and n are reduced by Δ:

$$x_M^0 = \frac{(m-\Delta)\,\varepsilon}{aw\,(1+n-\Delta)} - \frac{s_M}{aw}.$$

This level is smaller than the level given by (21) if and only if:

$$\begin{aligned}
\frac{m\varepsilon}{aw\,(1+n)} - 2m\frac{1+l}{aw\,(1+n)^2}\mu &> \frac{(m-\Delta)\,\varepsilon}{aw\,(1+n-\Delta)} \Rightarrow \\
\varepsilon\left[\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right] &> 2\mu\,(1+l)\,\frac{m}{(1+n)^2} \Rightarrow \\
\varepsilon &> 2\mu\,(1+l)\,\frac{\frac{m}{(1+n)^2}}{\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}} \\
&= 2\mu\,(1+l)\,\frac{\frac{1+n-\Delta}{1+n}m}{\Delta\,(1+l)} = 2\mu\frac{m\,(1+n-\Delta)}{\Delta\,(1+n)}.
\end{aligned}$$

When $\mu > 0$, we can write this as $\varepsilon/\mu > \overrightarrow{r}_M$ and, in this case, P's presence, when it induces institutional change, increases x_M, reduces x_L, and the total harm increases (the last two effects follow from the proof of Proposition 3).

When $\mu < 0$, we can write the inequality as $\varepsilon/\mu < \overrightarrow{r}_M$, so, when $\varepsilon/\mu > \overrightarrow{r}_M$, P's presence, when it induces institutional change, decreases x_M, increases x_L, and, again, the total harm increases (in line with the proof of Proposition 3).
(ii) Suppose K is small. If P is not present, K prefers decentralization, and x_M follows from (21):

$$x_M^0 = \frac{m\varepsilon}{aw\,(1+n)} - \frac{s_M}{aw}.$$

This level is smaller than the level given by (21) where, after an induced merger, m and n are reduced by Δ, if and only if:

$$\begin{aligned}
\frac{m\varepsilon}{aw\,(1+n)} &< \frac{(m-\Delta)\,\varepsilon}{aw\,(1+n-\Delta)} - 2\,(m-\Delta)\,\frac{1+l}{aw\,(1+n-\Delta)^2}\mu \Rightarrow \\
\varepsilon\left[\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}\right] &< -2\mu\,(1+l)\,\frac{m-\Delta}{(1+n-\Delta)^2} \Rightarrow \\
\varepsilon &< -2\mu\,(1+l)\,\frac{\frac{m-\Delta}{(1+n-\Delta)^2}}{\frac{m}{1+n} - \frac{m-\Delta}{1+n-\Delta}} \Rightarrow \\
\varepsilon &< \overleftarrow{r}_M \equiv -2\mu\frac{(1+n)\,(m-\Delta)}{\Delta\,(1+n-\Delta)}.
\end{aligned}$$

When $\mu > 0$, we can write this as $\varepsilon/\mu < \overleftarrow{r}_M$ and, in this case, P's presence, when it induces institutional change, increases x_M, reduces x_L, and the total harm increases (the last two effects follow from the proof of Proposition 3).

When $\mu < 0$, we can write the inequality as $\varepsilon/\mu > \overleftarrow{r}_M$, so, when $\varepsilon/\mu < \overleftarrow{r}_M$, P's presence, when it induces institutional change, decreases x_M, increases x_L, and, again, the total harm increases (in line with the proof of Proposition 3). *QED*

References

[1] Agrawal, A., Chhatre, A., and Hardin, R. "Changing Governance of the World's Forests." *Science*, Vol. 320 (2008), pp. 1460–62.

[2] Amacher, G. S., Koskela, E., and Ollikainen, M. "Royalty Reform and Illegal Reporting of Harvest Volumes Under Alternative Penalty Schemes." *Environmental and Resource Economics*, Vol. 38 (2007), pp. 189–211.

[3] Amacher, G.S., Ollikainen, M., and Koskela, E. "Corruption and Forest Concessions." *Journal of Environmental Economics and Management*, Vol. 63 (2012), pp. 92–104.

[4] Angelsen, A. and Kaimowitz, D. "Rethinking the Causes of the Models Deforestation: Lessons from Economic Models." *World Bank Research Observer*, Vol. 14 (1999), pp. 73–98.

[5] Ashenfelter, O. and Hosken, D. "The Effect of Mergers on Consumer Prices: Evidence from Five Mergers on the Enforcement Margin." *Journal of Law and Economics*, Vol. 53 (2010), pp. 417–66.

[6] Ashenfelter, O., Hosken, D., and Weinberg, M.C. "The Price Effects of a Large Merger of Manufacturers: A Case Study of Maytag-Whirlpool." *American Economic Journal: Economic Policy*, Vol. 5 (2013), pp. 239–61.

[7] Asker, J. and Heski B. "Raising Retailers' Profits: on Vertical Practices and the Exclusion of Rivals." *American Economic Review*, Vol. 104 (2014), pp. 672–86.

[8] Barro, J.R. and Cutler, D. "Consolidation in the Medical Care Marketplace: A Case Study from Massachusetts." *National Bureau of Economic Research,* No. w5957, 1997.

[9] Bayrak, M. and Marafa, L. "Ten Years of REDD+: A Critical Review of the Impact of REDD+ on Forest-dependent Communities." *Sustainability,* Vol. 8 (2016), pp. 1–22.

[10] Berkes, F. "Devolution of Environment and Resources Governance: Trends and Future." *Environmental Conservation*, Vol. 37 (2010), pp. 489–500.

[11] Burgess, R., Costa, F. J. M. and Olken B.A. "The Brazilian Amazon's Double Reversal of Fortune." MIT, `http://economics.mit.edu/files/12732`, 2019.

[12] Burgess, R., Hansen, M., Olken, B. A., Potapov, P., and Sieber, S. "The Political Economy of Deforestation in the Tropics." *Quarterly Journal of Economics*, Vol. 127 (2012), pp. 1707–54.

[13] Che, Y.-K. and Spier, K. E. "Exploiting Plaintiffs Through Settlement: Divide and Conquer." *Journal of Institutional and Theoretical Economics*, Vol. 164 (2008), pp. 4–23.

[14] Chiroleu-Assouline, M., Poudou, J.C., and Roussel, S. "North/South Contractual Design through the REDD+ Scheme." *FEEM Working Paper, No. 89,* 2012.

[15] Collier, P., and Venables, A. J. "Closing Coal: Economic and Moral Incentives." *Oxford Review of Economic Policy,* Vol. 30 (2014), pp. 492–512.

[16] Delacote, P. and Simonet, G. "Readiness and Avoided Deforestation Policies: on the use of the REDD Fund." *Working Paper 1312,* Chaire Economie du Climat, 2013.

[17] Dequiedt, V. and Martimort, D. "Vertical Contracting with Informational Opportunism." *American Economic Review*, Vol. 105 (2015), pp. 2141–82.

[18] Dranove, D., Rothman, D. and Toniatti, D. "Up or Down? The Price Effects of Mergers of Intermediaries." *Antitrust Law Journal,* Vol. 82 (2019), pp. 643–77.

[19] Gaudet, G. and Salant, S. W. "Increasing the Profits of a Subset of Firms in Oligopoly Models with Strategic Substitutes." *American Economic Review*, Vol. 81 (1991), pp. 658–65.

[20] Genicot, G. and Ray, D. "Contracts and Externalities: How Things Fall Apart." *Journal of Economic Theory*, Vol. 131 (2006), pp. 71–100.

[21] Gjertsen, H., Groves, T., Miller, D. A., Niesten, E., Squires, D., and Watson, J. "Conservation Agreements: Relational Contracts with Endogenous Monitoring." Forthcoming, *The Journal of Law, Economics, and Organization.*

[22] Goncalves, M. P., Panjer, M., Greenberg, T. S., and Magrath, W. B. *Justice for Forests: Improving Criminal Justice Efforts to Combat Illegal Logging*, World Bank, 2012.

[23] Harstad, B. "Buy Coal! A Case for Supply-Side Environmental Policy." Journal of Political Economy Vol. 120 (2012), pp. 77–115.

[24] Harstad, B. and Mideksa, T. K. "Conservation Contracts and Political Regimes." *Review of Economic Studies,* Vol. 84 (2017), pp. 1708–34.

[25] Hastings, J. "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California." *American Economic Review*, Vol. 94 (2004), pp. 317–28.

[26] Horn, H. and Wolinsky, A. "Bilateral Monopolies and Incentives for Merger." *RAND Journal of Economics*, Vol. 19 (1988), pp. 408–19.

[27] Hosken, D., Olson, L.M. and Smith, L. K. "Do Retail Mergers Affect Competition? Evidence from Grocery Retailing." *Journal of Economics & Management Strategy,* Vol. 27 (2018), pp. 3–22.

[28] Hosken, D. and Tenn, S. "Horizontal Merger Analysis in Retail Markets." Emek Basker ed., *Handbook on the Economics of Retailing and Distribution*, 2016, pp. 250–286.Edward Elgar Publishing, Cheltenham UK.

[29] Houde, J. "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline." *American Economic Review*, Vol. 102 (2012), pp. 2147–82.

[30] Inderst, R. and Shaffer, G. "Retail Mergers, Buyer Power and Product Variety." *Economic Journal*, Vol. 117 (2007), pp. 45–67.

[31] Inderst, R. and Wey, C. "Bargaining, Mergers, and Technology Choice in Bilaterally Oligopolistic Industries." *RAND Journal of Economics*, Vol. 34 (2003), pp. 1–19.

[32] Irawan, S. and Tacconi, L. "Reducing Emissions from Deforestation and Forest Degradation (REDD) and Decentralized Forest Management." *International Forestry Review*, Vol. 11 (2009), pp. 427–38.

[33] Kamien, M. I. and Zang, I. "The Limits of Monopolization through Acquisition." *Quarterly Journal of Economics*, Vol. 105 (1990), pp. 465–99.

[34] Landeo, C. M. and Spier, K. E. "Naked Exclusion: An Experimental Study of Contracts with Externalities." *American Economic Review*, Vol. 99 (2009), pp. 1850–77.

[35] Larson, A. M. and Soto, F. "Decentralization of Natural Resource Governance Regimes." *Annual Review of Environment and Resources*, Vol. 33 (2008), pp. 213–39.

[36] Marshall, R. C., and Marx, L. M. *The Economics of Collusion, Cartels and Bidding Rings.* 2012. MIT Press.

[37] Mason, C. F. "Optimal Contracts for Discouraging Deforestation with Risk Averse Agents." Mimeo. 2015.

[38] Mason, C. F. and Plantinga, A. J. "The Additionality Problem with Offsets: Optimal Contracts for Carbon Sequestration in Forests." *Journal of Environmental Economics and Management*, Vol. 66 (2013), pp. 1–14.

[39] McAfee, R. P. and Schwartz, M. "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity." *American Economic Review*, Vol. 84 (1994), pp. 210–30.

[40] Mendelsohn, R. "Property Rights and Tropical Deforestation." *Oxford Economic Papers*, Vol. 46 (1994), pp. 750–56.

[41] Milliou, C., and Petrakis, E. "Upstream Horizontal Mergers, Vertical Contracts, and Bargaining." *International Journal of Industrial Organization*, Vol. 25 (2007), 963–87.

[42] Oates, Wallace E. *Fiscal Federalism.* New York: Harcourt Brace, (1972).

[43] Perry, M. K. and Porter, R. H. "Oligopoly and the Incentive for Horizontal Merger." *American Economic Review*, Vol. 75 (1985), pp. 219–27.

[44] Phelps, J., Webb, E. E. L., and Agrawal, A. "Does REDD+ Threaten to Recentralize Forest Governance." *Science*, Vol. 328 (2010), pp. 312–13.

[45] Ribot, J.C., Agrawal, A., and Larson, A.M. "Recentralizing While Decentralizing: How National Governments Reappropriate Forest Resources." *World Development*, Vol. 34 (2006), pp. 1864–86.

[46] Salant, S. W., Switzer, S., and Reynolds, R. J. "Losses from Horizontal Merger: the Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium." *Quarterly Journal of Economics*, Vol. 98 (1983), pp. 185–99.

[47] Segal, I. "Contracting with Externalities." *Quarterly Journal of Economics*, Vol. 114 (1999), pp. 337–88.

[48] Segal, I. "Coordination and Discrimination in Contracting with Externalities: Divide and Conquer?" *Journal of Economic Theory*, Vol. 113 (2003), pp. 147–81.

[49] Silvia, L. and Taylor, C.T. "Petroleum Mergers and Competition in the Northeast United States." *International Journal of the Economics of Business*, Vol. 20 (2013), pp. 97–124.

[50] Steckel, J. C., Edenhofer, O. and Jakob, M. "Drivers for the renaissance of coal." *Proceedings of the National Academy of Sciences,* Vol. 29 (2015), pp. E3775–E3781.

[51] The Competition and Markets Authority. *Anticipated Merger Between J Sainsbury PLC and Asda Group Ltd: Final report*, the United Kingdom, 2019. https://assets.publishing.service.gov.uk/media/55ad1bdbed915d151e000011/Sainsburys_CGL.pdf