NUPI

[661] Paper

Uniform or Different Policies

Bård Harstad

No. 661 – 2004

Norwegian Institute of International Affairs
Norsk Utenrikspolitisk Institutt

Utgiver: NUPI

ISSN: 0800 - 0018

Besøksadresse: C.J. Hambrosplass 2d
Addresse: Postboks 8159 Dep.
0033 Oslo
Internett: www.nupi.no
E-post: pub@nupi.no
Fax: [+ 47] 22 36 21 82
Tel: [+ 47] 22 99 40 00

Uniform or Different Policies*

Bård Harstad

* I am indebted to Torsten Persson for his continuous support throughout this work. I have also benefited from the help of Philippe Aghion, Alberto Alesina, Patrick Bolton, Oliver Hart, Bob Staiger, Lucy White, seminar participants at IIES, Harvard University, the Norwegian School of Economics and Business Administration, University of Oslo, CICERO, the EEA Summer Meeting (2002), the North American Econometric Society Meeting (2002) and the 24th Norwegian Conference in Economics. Thanks also to Christina Lönnblad for editorial assistance. Financial support from Jan Wallander's and Tom Hedelius' Foundation and the Norwegian Research Council is gratefully acknowledged.

Key words: International negotiations, policy harmonization, side payments
JEL classification: C78, H77, D82, D72

[Abstract] I analyse the negotiation between two countries, or regions, that are trying to make an agreement in order to internalize externalities. Local preferences are local information, but reluctance to participate in the agreement is signaled by delay. Conditions are derived for when it is efficient to restrict the attention to policies that are uniform across regions - with and without side payments - and when it is optimal to forbid side payments in the negotiations. While policy differentiation and side payments let the policy be tailed to local conditions, they create conflicts between the regions and thus delay. If political centralization implies uniformity, as is frequently assumed in the federalism literature, the results describe when centralization outperforms decentralized cooperation. But the results also provide a foundation for this uniformity assumption and characterize when it is likely to hold.

1. Introduction

To get a more favorable deal, it is in any negotiator's interest to demonstrate a willingness to walk away. When bargaining power is private information, this motivates the parties to signal their strength and screen their opponent, even if this is costly. Such inefficiencies are inherent in any bargaining under private information.[1] But how can these inefficiencies be modified? This paper studies simple constraints on the bargaining agenda and characterizes when these are valuable.[2] By nailing the analysis to a particular example, the results shed light on existing institutions as well as controversies in the literature on fiscal federalism.

I study two regions trying to coordinate their policies to internalize externalities. By reducing or cleaning pollution, say, each region does not only improve the air quality in its own region; but also in the other region. Local values of clean air, however, are local knowledge. The regions negotiate the agreement in an alternating offer bargaining game, where each region can delay as long as it desires before making an offer.[3] While this bargaining game should be interpreted as an example only, it is useful in that (i) it implements the best "reasonable" mechanism, and (ii) it provides intuition for how this is achieved. In this context, I ask the following three questions. When would the regions, behind a veil of ignorance, prefer to commit to "harmonized agreements" where contributions must be uniform across regions? How does the possibility to use side payments affect the answer? Should side payments be possible?

Economists typically ignore these questions, presuming that policy differentiation and side payments obviously are good because they increase the flexibility in negotiations. But these questions are important for policy advisors. Federal politics are quite often characterized by extensive harmonization. In the European Union, for example, Article 100a (Single European Act) calls for approximation of laws and harmonization measures, which has led to an explosion of directives calling for uniform policies for waste, chemicals, and other measures (discussed by McCormick, 2001). Such clauses appear puzzling. While economists recognize the need to coordinate policies when there are externalities, the first-best policy ought to reflect local conditions. By allowing our two regions negotiate unconstrained, the region with the highest value of clean air must clean most in equilibrium. Although this outcome is due to the difference in bargaining power, it is beneficial that most cleaning is located where it is mostly appreciated. But negotiations may be costly when there is private information. By giving in early, a region reveals a high willingness to pay for the agreement, and it must thus bear the lion's share of the cost. Thus, to get a more favorably deal, the regions may want to signal reluctance to

[1]For a survey, see Ausubel, Cramton, and Deneckere (2002).

[2]The mechanism design approach to this problem, initiated by Myerson and Sattertwhite (1983) excludes unreasonable outcomes and picks the best alternative in the remaining set. Although useful in providing upper bounds to what is achievable, the best mechanism typically hinges on the players' beliefs in complex ways, and it requires considerable commitment. The alternative approach, pursued in this paper, is to explicitly analyze negotiations and investigate how the outcome depends on the bargaining agenda.

[3]In this respect, the bargaining game is similar to those suggested by Admati and Perry (1987) and Cramton (1992), as is the equilibrium I consider.

participate by strategic delay.[4] A commitment to uniform policies, on the other hand, reduces the conflict of interest between the regions. Since the regions will have to make equal contributions to an agreement in any case, they do not find it worthwhile to signal bargaining power by delay. To judge whether the policy should be differentiated instead of uniform across regions, the gains from differentiation should be compared to the increased cost of delay. Proposition 1 shows how the optimal bargaining agenda depends on the expected value of an agreement, the amount of externality and heterogeneity.

While the above case is a nice benchmark case, it is typically presumed that allowing for side payments increases efficiency. If the regions are closely integrated, issue linkages and logrolling will be intrinsic in the political debate. It is then realistic to let the regions negotiate over side payments in addition to the policy. The policy will then typically be optimally differentiated in equilibrium, since the region contributing less can compensate the one contributing more. To exploit these gains from trade, it is always better to allow policy differentiation (Proposition 2). However, Proposition 3 describes when it is even better to prohibit both side payments and differentiation, as this reduces the the conflict of interests, and thus delay.

Bargaining agendas:

		Differentiation?	
		no	yes
Side payments?	no	x	d, x
	yes	s, x	d, s, x

Are side payments good? Having analyzed the outcome with and without side payments (by horizontal comparisons in the above table), it is constructive to compare the two cases (by a vertical comparison) to examine the isolated impact of side payments. Side payments can typically be facilitated by logrolling, issue linkages and giving the negotiators discretion over several political issues. Whether it is desirable to introduce side payments in this way is thus an important question.[5] By allowing side payments, any action increasing total utility can be implemented as a Pareto improvement by making the winner compensate the loser. It is therefore a common presumption that *side payments are needed to reach the best result* (Cesar and de Zeeuw, 1996, p. 158). The existing literature on issue linkages (see e.g. Inderst, 2000) emphasizes such gains from trade indeed, but it also detects distributive effects. The present paper contributes to this debate. On the one hand, another rationale for side payments is presented. When

[4] This is related to "why stabilization is delayed". Alesina and Drazen (1991) study a war of attrition between legislators trying to stabilize the economy. The proposal-maker must bear the lion's share of the stabilization cost. Hsieh (2000) endogenizes this assumption: accepting early stabilization reveals a politician's willingness to pay. Another politician can therefore safely require that the first bears most of the stabilization cost. Anticipating this, every politician is reluctant to propose stabilization, and stabilization is delayed.

[5] Negotiations in the European Council (consisting of the heads of states who have a lot of discretion over alternative political issues) are likely to involve issue linkages and thus side payments. Side payments are less likely to be possible for issues determined by the Council of Ministers, where the various ministers lack discretion over other policies. Since the European Union's decision-making procedures are currently under debate, and since countries generally disagree on the extent of issue linkages, it is both important and timely to ask whether side payments should be desired.

the regions negotiate over side payments as well as the policy, a region can signal its type by the combination it suggests. Signaling by the proposed direction of trade can thus replace costly delay as a signaling device. On the other hand, the distributive effect of side payments may increase the conflict of interest between regions. Without side payments, a reluctant region is likely to contribute less to the agreement; with side payments, it may even be able to acquire transfers from the other region. If so, bargaining power pays off more, and the incentive to signal reluctance increases when side payments are possible. For side payments to be beneficial, the gains from trade must be larger that the increased cost of delay. Depending on the expected value of the agreement, the externality and the heterogeneity, Proposition 4 presents conditions under which introducing side payments is actually bad.

The results of this paper describe the *optimal* bargaining agenda. In the model, policy differentiation and side payments will always be included unless regions are able to *commit* to the optimal agenda in advance (before they know their own types). The necessity to commit ties this paper to the literature on fiscal federalism (surveyed by Oates, 1999) in two interesting ways. To some extent, frequent interaction facilitates the possibilities to commit, since sticking to the optimal agenda today is motivated by cooperation tomorrow. In other cases, it is necessary to commit by writing a more formal constitution. For either reason, regions in a federal union should be better able to commit to uniform policies when this is optimal. Thus, the theory predicts more uniform policies within than across federal unions. This is exactly the critical "uniformity assumption" made in the federalism literature,[6] which often is criticized for being ad hoc.[7] The analysis in this paper provides a microeconomic foundation for the uniformity assumption, and characterizes when it is likely to hold.

Suppose that we rely on the assumption that centralization implies uniform policies. Centralizing the policy is then a certain way of committing to uniform policies. The comparison between uniform and differentiated policies becomes identical to the comparison between centralization (requiring uniform policies) and decentralization (where the

[6] Oates (1972) initiated an entire literature based on this assumption. Recently, Alesina and Spoalore (1997) have analyzed the optimal and equilibrium size of nations when the benefit of a large size is increasing returns to scale, while the cost is that the policy must be uniform across heterogeneous regions. Bolton and Roland (1997) investigate the breakup of nations under similar assumptions, while Ellingsen (1998) studies political integration. In his survey over the fiscal federalism literature, Oates (1999, p. 1130) states that *There is clearly some kind of trade-off here between internalizing spillover benefits (and costs) and allowing local differentiation.*

[7] Lockwood (2002) claims that the uniformity assumption is *not derived from any explicit model of government behavior*, and Besley and Coate (2003) therefore declare that *Oates' analysis is suspect.* There are some papers, however, suggesting how the political game may induce uniformity. Cremer and Palfrey (2000) show how a majority might vote for a federal mandate, i.e. a minimum level for local policies, which might be too strict. The reason is that voters unaffected by the mandate (sincerely) vote together with those benefiting from a stricter mandate. Panagariya and Rodrik (1993) argue that a uniform tariff across industries might be optimal to reduce lobbying and tie the hands of politicians favoring certain groups. By similar arguments, Besharov (2002) shows that uniform policies may be optimal to avoid costly lobbying. It is less clear why uniformity should result from asymmetric information. In a situation with private information, Bordignon, Manasse and Tabellini (2001) show that it is still optimal to differentiate regional transfers even if this requires that the region with the lowest tax base must signal this by an inefficiently large tax rate.

negotiating regions are free to propose differentiation and perhaps side payments). Thus, Proposition 1-4 may be interpreted as conditions for when centralization outperforms decentralization.[8] Traditionally, the fiscal federalism literature evaluates decentralization vs. centralization assuming that if the policy is decentralized, there will be no political coordination. However, even if the policy is decentralized, the regions certainly have incentives to negotiate whenever externalities exist. It is therefore reasonable to relax this assumption, as is done in this paper. While some of the traditional insights are confirmed, others are turned upside-down.

The paper is organized as follows. Next section models the economy and the bargaining game in a context without side payments. The following section describes the equilibrium and its properties, and discusses circumstances under which a uniform policy is better than a differentiated one. Section 4 repeats this exercise with side payments, while Section 5 compares the two cases and derives conditions under which side payments should be prohibited. The alternative interpretations of the results are discussed in the concluding section. All proofs are found in the Appendix.

2. The Model

2.1. The Environment

Two regions, A and B, consider whether to undertake a joint public project and, if so, how its total cost of 1 should be distributed between them. To fix ideas, suppose A and B suffer from a symmetric regional environmental problem. A fraction e of A's emission crosses the border and pollutes region B, while a fraction $1-e$ remains as local pollution in region A. A's value of clean air is v_A, but there is a unit cost to clean or reduce emission. Let x_A denote the amount A cleans or reduces emission. Making symmetric assumptions for region B, the regional (von Neumann-Morgenstern) utilities are

$$\begin{aligned} u_A &= [(1-e)x_A + ex_B]\, v_A - x_A \\ u_B &= [(1-e)x_B + ex_A]\, v_B - x_B. \end{aligned} \tag{2.1}$$

A region pays the entire cost of its cleaning, while it only receives a fraction $(1-e)$ of the benefit. Since the other region enjoys the remaining fraction e, each region prefers the other region to clean as much as possible. If $e = 1/2$, cleaning is a pure public good.

Assume that a region will clean nothing unless an agreement is formed, because $v_i(1-e) < 1$.[9] so that it is not in the interest of one region to clean alone. Relative to this default outcome, the regions are considering a joint public project that will reduce total emission by one unit, i.e. $x_A + x_B = 1$. x_A is therefore A's contribution to the public project. A *uniform policy* requires that both regions have the same environmental standard, or that they both reduce emission by the same amount. In either case, this implies that $x_A = x_B$.

[8] This tradeoff is quite similar to that analyzed by Bolton and Farrell (1990). They study firms' entry into a market. While the cost of decentralizing this decision might be delay (as well as duplication when both firms enter), the benefit is that the most efficient firm is likely to enter first. A clumsy government, in contrast, will immediately but randomly select one firm.

[9] In statements that may be true for either region, I let i denote any of these, i.e. $i \in \{A, B\}$.

Assume that $v_i > 1$, such that both regions benefit from a uniform agreement relative to no agreement.[10] If the policy is not uniform, the amount of *differentiation* is defined as

$$d \equiv x_B - x_A.$$

There are potential benefits from differentiating the policy if the regions are heterogeneous ($v_A \neq v_B$). This can be seen from rewriting (2.1) as

$$\begin{aligned} u_A &= \frac{1}{2}[v_A - 1] + \frac{1}{2}[1 - (1-2e)v_A]d \qquad (2.2) \\ u_B &= \frac{1}{2}[v_B - 1] - \frac{1}{2}[1 - (1-2e)v_B]d \end{aligned}$$

and defining total welfare as

$$u \equiv u_A + u_B = \frac{1}{2}(v_A + v_B) - 1 + \left(\frac{1}{2} - e\right)(v_B - v_A)\,d. \qquad (2.3)$$

If $e < 1/2$, pollution is mainly a local problem, and $d > 0$ is efficient if and only if $v_B > v_A$: it is efficient that the region with the highest value of clean air reduces its emission most. If $e = 1/2$, however, cleaning is a pure public good and there is no value from differentiating the policy. If $e > 1/2$ (e.g. due to the plants' strategic location), most of the emission crosses the border, and $d < 0$ is efficient if and only if $v_B > v_A$: it is optimal that the region with the lowest value of clean air reduces its emission most.

However, A and B have conflicting interests in *how* the policy should be differentiated. Each region prefers that the other region contributes most, and the regional contributions are determined by the negotiations. These negotiations are complex since local preferences are local information.[11] Each region i knows only its own type $v_i \in \{\underline{v}, \overline{v}\}$, and the fact that the other region's type is either $\underline{v}$ or $\overline{v} > \underline{v}$ with equal probability.[12]

2.2. A Bargaining Model

In the above environment, regions A and B try to make an agreement. This subsection suggests a bargaining game describing their negotiations. Naturally, this bargaining game

[10] I thus abstract from the issue of participation. If $v_i < 1$ were possible, differentiation or side payments would be necessary to encourage i to participate, which would certainly reduce the case for uniform policies (see e.g. Hoel, 1992).

[11] This is a standard assumption in the fiscal federalism literature, and it is also empirically plausible. For example, in a discussion of European environmental policies, Mäler (1991) observes that *the control costs and environmental damage in one country are known to that country only.*

[12] The model can be interpreted and modified in several alternative ways. First, instead of negotiating the allocation of costs, regions may negotiate how to share a cake. The utility function above can be rewritten as $u_A = (1 + ev_A)\,x_B + (1-e)v_A x_A - 1$. Let the regional cost of the cake be fixed and equal to one for each region; x_B be A's share of the cake; x_A be B's share. The externality (from cake consumption) should now be interpreted as $1 - e$, instead of e as before. Otherwise, the results will be the same.

Second, instead of heterogeneity related to values, the heterogeneity might be related to costs. This requires me to slightly rewrite the model, although the analysis and the trade-offs would be similar.

Third, allowing for observed heterogeneity in addition to the unobservable heterogeneity above is straightforward. A's type could either be $\underline{v}_A$ or $\overline{v}_A$, while B's type could be $\underline{v}_B$ or $\overline{v}_B$. Were the observable heterogeneity larger, a differentiated policy would be more likely to be better.

should only be interpreted as an example since the regions may negotiate differently. Nevertheless, the following procedure is useful because (i) it implements the most efficient "stable" and "robust" mechanism (defined in the next section), and (ii) it provides intuition for how implementation is achieved.

The bargaining game is quite standard. The regions make alternating offers over d, A makes the first offer, time is continuous and the time horizon is infinite. An early agreement is preferred to a later one, since i's present value of an agreement settled at time t is $\delta^t u_A$, where $\delta < 1$ is the regions' common discount factor. The *minimum* time between offers is arbitrarily small (and approaching zero). However, I relax the standard assumption that a region *must* make a proposal at a certain time. Each region is allowed to delay as long as it wishes before making an offer. This provides a way for the regions to signal their types. In this respect, the game is similar to the seller-buyer bargaining games proposed by Admati and Perry (1987) and Cramton (1992), as is the equilibrium. Admati and Perry find a unique sequential equilibrium when the seller's type is known, while the buyer's type is either high or low. Cramton describes an equilibrium in a symmetric game with two-sided private information, and where the distribution of types is continuous.[13]

3. Uniform or Different Policies?

The first subsection describes a sequential equilibrium outcome when the policy can be differentiated. Although the equilibrium is not unique, the next subsection presents some of its attractive features, which justify why I let this outcome represent the case with differentiation. The third subsection compares this outcome to the outcome when policies must be uniform, and derives Proposition 1.

3.1. The Outcome with Differentiation

Suppose that both regions' types were common knowledge. The above bargaining game would then have the same unique equilibrium as in the standard Rubinstein (1982) bargaining game. The amount of differentiation would be given by d_R, defined as

d_R: B's type (3.1)

A's type	$\underline{v}$	$\overline{v}$
$\underline{v}$	0	d'
$\overline{v}$	$-d'$	0

[13]In this paper, as in most of the literature on bargaining with private information, bargaining power is signaled by costly delay. Alternatively, bargaining power may be signaled by proposing a suboptimal or an incomplete agreement. In fact, all results in this paper continue to hold if, instead of delaying to t, each region can credibly reduce the total amount of cleaning in the relevant agreement from 1 to δ^t. Instead of delaying, a low-type region would then signal its bargaining power by proposing a less ambitious project.

where[14]

$$d' \equiv \frac{1}{2}\left[\frac{\overline{v}-1}{1-\overline{v}(1-2e)}-\frac{\underline{v}-1}{1-\underline{v}(1-2e)}\right]. \tag{3.2}$$

If B has a high value of clean air, B is very eager to settle the agreement quickly. Since eagerness reduces the bargaining power, a low-type A forces B to contribute most to the agreement, so then $d = d' > 0$.

When local preferences are local information, the final agreement will still be the one determined by (3.1), but only after each low-type region has credibly signaled its type by a sufficient delay. The outcome will be the following. Suppose A is of high type. Then, A immediately proposes that the two regions should make equal contributions ($d = 0$). A high-type B immediately accepts. A low-type B, however, rejects A's offer and delays until time t_1 before suggesting (by proposing $-d'$) that A contributes most. This is immediately accepted by A. Suppose instead A is of low type. Then, A does not make any immediate offer. Instead, A delays until t_1 before suggesting (by proposing d') that B contributes most. A high-type B immediately accepts. A low-type B, however, rejects A's offer and delays until t_2 before suggesting that they make equal contributions ($d = 0$), which A immediately accepts.

When a region accepts an offer d_R, it does so because it is convinced that the other region is of a certain type. Each delay is exactly sufficiently long to credibly signal that the region is of a low type. A high-type region is less patient, and cannot afford such a delay. The low-type region therefore separates itself from the high-type, by taking an action (delay) that the other type cannot afford.[15] This requires that

delay: (3.3)

A's type \ *B*'s type	$\underline{v}$	$\overline{v}$
$\underline{v}$	t_2	t_1
$\overline{v}$	t_1	0

[14] Note that an affine transformation of the utilities gives

$$\widetilde{u}_A \equiv \frac{u_A}{\frac{1}{2}[1-v_A(1-2e)]} = w_A + d$$
$$\widetilde{u}_B \equiv \frac{u_B}{\frac{1}{2}[1-v_B(1-2e)]} = w_B - d,$$

where

$$w_i \equiv \frac{v_i - 1}{1-v_i(1-2e)}$$

is region i's willingness to pay for the agreement in terms of d. In the Rubinstein (1982) alternating offer bargaining game, as the time between offers approaches zero, d will be set such that $\widetilde{u}_A$ and $\widetilde{u}_B$ are equalized:

$$d_R = \frac{w_B - w_A}{2},$$

which gives (3.1). This will still be the equilibrium when regions have the possibility to delay the agreement, since no region could benefit from a delay.

[15] This is possible since the utility function $\delta^t u_i$ fulfills the single-crossing property.

where

$$\delta^{t_1} = 1 - \left[\frac{(1-\overline{v}(1-2e))d'}{\overline{v}-1+(1-\overline{v}(1-2e))d'}\right] \tag{3.4}$$
$$\delta^{t_2} = 1 - 2\left[\frac{(1-\overline{v}(1-2e))d'}{\overline{v}-1+(1-\overline{v}(1-2e))d'}\right].$$

3.2. Equilibrium Properties

While the previous subsection merely described the equilibrium outcome, this subsection characterizes some of its properties and justifies the attention it will be paid.

As stated by the following Lemma 1, the above outcome can be supported as a sequential equilibrium. Moreover, it is the unique equilibrium if we apply a certain *optimistic intuitive criterion* to refine the set of equilibria. No pooling equilibria then exist.[16] In addition, the above outcome has several attractive properties. It is symmetric, and there is no first-mover advantage. Since only low-type regions delay, an agreement is settled earlier if it is more valuable. And when the regions finally settle the agreement, they do so at "fair" terms, i.e. according to the Rubinstein outcome (3.1). This feature is of particular importance in our context. It is widely agreed that international negotiations must be self-enforcing, since external enforcement mechanisms are seldom available. After an agreement has been formed, A can leave the agreement with the only consequence that the agreement breaks down.[17] If A does so, the two regions have incentives to renegotiate a new agreement. If, at this point in time, A's and B's types are revealed, they immediately agree on the Rubinstein bargaining outcome (3.1). Anticipating this, no other agreement than (3.1) is robust for such a unilateral request to renegotiate. Thus, it is reasonable to require $d = d_R$ to make self-enforcing agreements renegotiation-proof.

Definition 1: An agreement d is *stable* if and only if $d = d_R$.

To fully evaluate the above equilibrium outcome, however, we should ask what else could be achieved. Notwithstanding how the regions solve their problem, their method could be substituted by a mechanism where honest revelation is an equilibrium. In our context, we can define a mechanism in the following way:

Definition 2: Let $\widehat{v} = \{\underline{v}, \overline{v}\}$ be each player's strategy set. A *mechanism* is a mapping $M : \widehat{v}^2 \to \mathbb{R} \times \mathbb{R}^+$, which determines an outcome (d, t) for each pair of possible types the regions may announce.[18]

Unrestricted mechanism design is often criticized as requiring too much of the institutional environment. First, the optimal mechanism typically implies ex post suboptimal

[16] The definitions and discussions of these concepts are relegated to the Appendix, to separate the theoretically inclined reader from the more applied one.

[17] Barrett (2001) writes that *the rules for international law allow countries to withdraw from an international treaty, at least after giving sufficient notice; and, as to reaffirm this freedom, nearly all treaties include an explicit provision for withdrawal.*

[18] There is no loss of generality by letting the mechanism be deterministic. Instead of letting the mechanism specify a probability p for the agreement to be formed, it can simply let the agreement be formed at date t, where $\delta^t = p$.

outcomes. If, however, the agreement must be self-enforcing, as argued above, then the mechanism should be restricted to outcomes that are ex post stable. I will say that a mechanism M is *stable* if $d = d_R$ whenever the regions announce their true types.

Second, the optimal mechanism typically requires the regions to simultaneously announce their types, thereby preventing the behavior of one from depending on that of the other. This is not how regions solve problems in practice, however. In negotiations, as the one described above, regions reveal their types sequentially, and perhaps little by little. While a region might be uncertain about its opponent's value before making an offer, the opponent's value might just as well have been revealed. Moreover, in reality, it is hard to separate exactly what is private information, and what regions know about each other (perhaps by espionage). A mechanism *only* working whenever types are private information, does not seem very appealing. In contrast, I will define a robust mechanism to be one that works *even* if a region should be aware of the other region's type.

Definition 3: A mechanism M is *robust* if it is incentive compatible, whether or not one region knows the other region's type.[19]

If mechanisms must be stable and robust, then the regions can actually not do better than in the equilibrium outcome described above.

Lemma 1: (i) Equations (3.1)-(3.4) characterize a sequential equilibrium outcome, (ii) which is unique under the *optimistic intuitive criterion.* Moreover, (iii) it implements the most efficient stable robust mechanism.[20]

3.3. When is Uniformity Better?

Suppose that before knowing their own types, the two regions anticipate the above outcome. Since each region may be of low or high type with equal probability, their total expected utility can be written as

$$u^d = \frac{1}{4}(\overline{v}-1) + \frac{1}{2}\left[\frac{\overline{v}+\underline{v}}{2} - 1 + \left(\frac{1}{2} - e\right)(\overline{v}-\underline{v})\, d'\right]\delta^{t_1} + \frac{1}{4}(\underline{v}-1)\delta^{t_2}. \tag{3.5}$$

Suppose, further, that the two regions would be able to commit to uniform policies ($d = 0$) if they ever reached an agreement. Would they make such a commitment? If they did, the proceeding bargaining outcome would be simple. A would immediately suggest an agreement, and B would immediately accept, whatever are their types. There would be no conflict of interest, as they would not have discretion over d. Thus, there would be no point in signaling reluctance by delay, as there would be no way of exploiting bargaining

[19] This implies that a robust mechanism is incentive compatible for *any* beliefs that one region may have about the other region's type.

[20] How restrictive are the requirements that the mechanism must be stable and robust? By relaxing the first of these, and assuming that $e < 1/2$, the best separating mechanism implies that d' is as large as possible, while the delays are still given by (3.3). The following result (Proposition 1) is not altered. If relaxing the second requirement, the best mechanism dictates no delay if only one region announces low type, but a longer delay if both do. But if $e \geq 1/2$, no mechanism can do better than a uniform policy.

power.[21] The total expected utility would be

$$u^0 = \frac{1}{4}(\overline{v} - 1) + \frac{1}{2}\left(\frac{\overline{v} + \underline{v}}{2} - 1\right) + \frac{1}{4}(\underline{v} - 1). \tag{3.6}$$

By comparison, differentiation provides costs as well as benefits. The potential benefit is that the region with the highest value of clean air will reduce its emission most, the cost is that such an agreement only occurs with some delay. Such delays are necessary for a low-type region to credibly signal its type which, in turn, is necessary to make the other region accept a larger contribution.

Define the expected value of clean air as

$$v \equiv \frac{\underline{v} + \overline{v}}{2},$$

and the heterogeneity by the relative difference in the two types' net value of a uniform agreement:

$$h \equiv \frac{\overline{v} - 1}{\underline{v} - 1} > 1.$$

It turns out that a uniform policy is better whenever

$$h\left[2(v-1)\left(\frac{1-2e}{e}\right) - 1\right] \leq 3, \tag{3.7}$$

that is, when the externality is large while the heterogeneity and the expected value of the agreement are low. The basic intuition for this is the following. If the externality e is low, it is beneficial that the high-type region cleans most, since this will imply that the air is cleanest where this is most appreciated. Thus, the differentiation following from the bargaining game is valuable. If $e \approx 1/2$, however, cleaning is (almost) a pure public good and it is of no importance where it is located, since the cost is the same in both regions. The value of differentiation is then low. If $e > 1/2$, it would be optimal that the low-type region contributed most. In equilibrium, however, the high-type region contributes most, since it has the lowest bargaining power. Allowing for differentiation would then clearly be perverse. Thus, the benefit from a differentiated policy decreases when e becomes larger. The cost, it turns out, increases. As e increases, each region benefits more from the other region's contribution, and the high type becomes more tempted to imitate the low-type's strategy. To credibly signal bargaining power, delay must increase. In sum: if e increases, the cost of differentiation increases while the benefit decreases, and a uniform policy becomes better.

Suppose that $e \leq 1/2$. If v increases, there is an increase in the gains from cleaning at home. The value of convincing the other region to contribute more, $(1 - (1-2e)\, v_i)$, decreases. In particular, the high-type region becomes less tempted to delay for the only

[21] If utilities were concave in x_i, however, different types would prefer different levels of cleaning, even if the contributions were bound to be equal in both regions. In that case, it would still be a conflict of interest between the regions under uniform policies. If the utility functions were not extremely concave, however, the conflict of interest would be even larger if differentiation were allowed, which would increase the delay.

purpose of contributing less. Thus, delay decreases, and differentiation is more likely to be best. If $e > 1/2$, however, most of the domestic pollution comes from the other region. Then, it becomes more important that the other region does most of the cleaning, which makes it more tempting to signal bargaining power, and there is an increase in delay. This makes a uniform policy even more superior.

If the heterogeneity h increases, the *value* of differentiation increases directly, which makes a differentiated policy better relative to a uniform one.

It should be noticed that the discount factor δ affects neither the cost nor the benefit of differentiation. If δ decreases, delay becomes more costly, but there is a corresponding decrease in the delay required to credibly signal reluctance. The cost of delay remains the same. The benefit of differentiation comes closer in time, but its present value remains constant.

The above discussion is not complete, however. If e, h, or v changes, so does the amount of differentiation d'. And when d' changes, so do both the cost and the benefit of differentiation. If d' increases, for example, the amount of differentiation increases, and thus also the potential benefit. But a larger d' makes the high-type more tempted to imitate the low-type, and to credibly signal bargaining power, delay must increase. The proof of Proposition 1 shows that costs and benefits increase similarly when d' increases, and the two effects cancel.

In reality, d' may not be determined by negotiations alone. Economic or technological constraints may limit to what extent the policy can be differentiated, such that $d \in [-D, D]$ for some $D \geq 0$. If this constraint were binding, i.e. if $d' > D$, it is easily shown that the outcome (3.1)-(3.4) continues to describe the equilibrium if just d' is replaced by D. As argued above, the amount of differentiation (d' or D) does not affect whether a uniform policy is better. The following proposition holds in any case.[22,23]

Proposition 1: $u^0 \geq u^d$ if and only if condition (3.7) holds. This is more likely if the externality e is large, the heterogeneity h small, and the value v low.

Some policies call for more harmonization than others. For the European Union, for example, it is enlightening to compare Articles 100a and 130s in the Single European Act. While the latter Article applies to environmental issues in general, the former encourages harmonization measures particularly for policies affecting the internal market, where the externality is likely to be larger. Interestingly, derogation (policy differentiation) is not possible under Article 100a, while it is under Article 130s. Moreover, uniform policies are easier to implement under Article 100a, since this requires a qualified majority only, as opposed to the unanimity required by Article 130s. Both differences seem to be in line with Proposition 1.

[22] A careful reader may suggest that $D = 1$, since $d > 1$ would imply that A increases its emission by signing the agreement. But some types of policy are easier to differentiate than others, and since there exist contexts where both $D < 1$ and $D > 1$ might be reasonable, I do not specify a value for D.

[23] The above analysis is restricted to the comparison between zero differentiation and equilibrium level of differentiation. Could an interior solution be optimal, making D endogenous? The answer is no. If d' were replaced by D, it can be shown that there is some optimal value D^* maximizing u^d. However, unless (3.7) holds, $D^* > d'$, such that it would never be optimal to restrict D below d'.

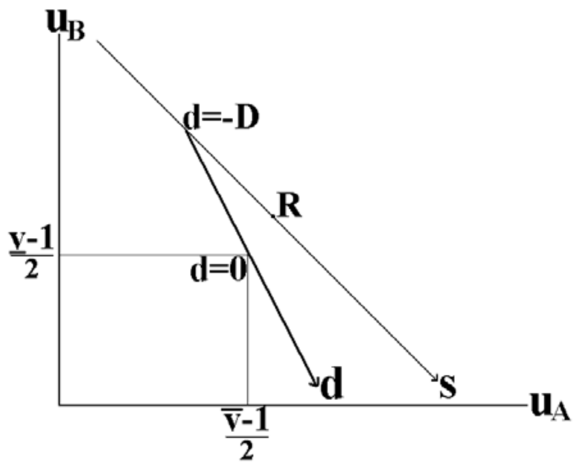

Figure 4.1: *The static Pareto frontier when the agenda is* (d, s), $e < 1/2$, *A is of high type and B is of low type. R indicates the outcome if information is symmetric.*

4. Introducing Side Payments

-*Side payments may not be relevant if we consider coordination in a single issue. But they may be highly relevant when the coordinating countries are integrated also in other areas of policy, as in Europe today* (Persson and Tabellini, 1995, p. 2000). As issue linkages and logrolling become intrinsic in the political debate, side payments can be included and perhaps not excluded from the bargaining agenda. Moreover, economists typically presume that side payments improve the efficiency of negotiations. For these reasons, I now introduce s as a (possible negative) side payment from B to A. Transaction costs related to such side payments are assumed to be negligible. We can then rewrite (2.2) as

$$\begin{aligned} u_A &= \frac{1}{2}[v_A - 1] + \frac{1}{2}[1 - v_A(1-2e)]d + s \\ u_B &= \frac{1}{2}[v_B - 1] - \frac{1}{2}[1 - v_B(1-2e)]d - s. \end{aligned} \tag{4.1}$$

The bargaining game is similar to that above, but now, each proposal is a pair (d, s). The static Pareto frontier is drawn in Figure 4.1.

4.1. The Outcome with Side Payments and Differentiation

If information were complete, the bargaining outcome would be an immediate agreement where $d \in [-D, D]$ would maximize the sum of the utilities while s would be set so as to equalize A's and B's utilities.[24] If $v_A = v_B$, the bargaining outcome (d, s) would be

[24] This can simply be shown by using arguments similar to Rubinstein (1982).

$(0,0)$.[25] Otherwise (d,s) are given by

	$(v_A,v_B)=(\underline{v},\overline{v})$	$(v_A,v_B)=(\overline{v},\underline{v})$
$e \le 1/2$	$(D,\underline{s})$	$(-D,-\underline{s})$
$e > 1/2$	$(-D,\overline{s})$	$(D,-\overline{s})$

(4.2)

where

$$\begin{aligned}\underline{s} &\equiv \frac{1}{4}(\overline{v}-\underline{v})-\frac{D}{4}[2-(\overline{v}+\underline{v})(1-2e)]\\ \overline{s} &\equiv \frac{1}{4}(\overline{v}-\underline{v})+\frac{D}{4}[2-(\overline{v}+\underline{v})(1-2e)].\end{aligned}$$

If the regional contributions were equal, side payments would go from the high-type region to the low-type region, since the former benefits more from an agreement than the latter. But if one region contributes more than the other, it must be compensated. The net side payment will consist of the sum of these two forces. If $e < 1/2$, most of the pollution is local and it is optimal that the high-type region cleans as much as possible. The two forces then pull in opposing directions, and it is unclear whether the side payment $\underline{s}$ that equalizes utilities is positive or negative. If $e > 1/2$, however, most of the emission crosses the border and it is optimal that the low type cleans most. The side payment to the low type is then $\overline{s} > \underline{s}$, and is clearly positive. If pollution is a pure public good ($e = 1/2$), it is of no importance where cleaning takes place, as long as the side payment equalizes utilities.

When local preferences are local information, the final agreement will still be that determined by (4.2), but only after each low-type region has credibly signaled its type. The outcome will be the following. Suppose $e \le 1/2$. If region A is of high type, it proposes $(d,s) = (0,0)$ at $t = 0$. A high-type B immediately accepts. A low-type B rejects A's offer and delays to t_1^s before it counteroffers $(-D,-\underline{s})$, which A accepts. If region A is of low type, it does not make any immediate offer. Instead, A delays to t_1^s before proposing $(D,\underline{s})$. A high-type B immediately accepts. A low-type B rejects A's offer and delays to t_2^s before it counteroffers $(0,0)$, which A immediately accepts. If $e > 1/2$, the game is similar, but now a low-type region suggests contributing *most* (against the compensation $\overline{s} > \underline{s}$), because this gives the largest total utility.

When a region accepts an offer, it does so because it is convinced that the other region is of a certain type. Each delay is exactly sufficiently long to credibly signal that the region is of a low type. A high-type region finds the low-type region's strategy unattractive, for two reasons. First, a high-type region is less patient, it cannot afford such a delay. Second, a low-type region *pays* the other region to contribute most (or least, if $e > 1/2$). A high-type region, in contrast, would benefit from the opposite agreement. The regions are exploiting "gains from trade" by allocating cleaning where it is most valuable. A region can thus signal its type by proposing a certain direction of trade.[26] If $D\,|1-2e|$ is large, the gains from optimal differentiation are large, the high-type is little tempted to imitate

[25] If regions were of the same type, and if transaction costs were identical to zero, the choice of d would be of no importance as long as s is such that the utilities are equal. A small but negligible transaction cost would make $s = d = 0$ the strictly better agreement, however.

[26] That a player can signal its type by the proposed direction of trade is related to the result by Cramton,

the low-type, and the necessary delay to separate the two types decreases. In fact, if $D\,|1-2e| \geq 1$, proposing a direction of trade is a sufficient signal: delay is not necessary and the bargaining outcome is first best. If $D\,|1-2e| < 1$, however, it is necessary with the following

delay: B's type

A's type		$\underline{v}$	$\overline{v}$
	$\underline{v}$	t_2^s	t_1^s
	$\overline{v}$	t_1^s	0

where

$$\begin{aligned} \delta^{t_1^s} &= 1-\left[\frac{(\overline{v}-\underline{v})\,(1-|1-2e|\,D)}{2(\overline{v}-1)+(\overline{v}-\underline{v})\,(1-|1-2e|\,D)}\right] \\ \delta^{t_2^s} &= 1-2\left[\frac{(\overline{v}-\underline{v})\,(1-|1-2e|\,D)}{2(\overline{v}-1)+(\overline{v}-\underline{v})\,(1-|1-2e|\,D)}\right]. \end{aligned} \tag{4.3}$$

4.2. Equilibrium Properties

As stated by Lemma 2, the above outcome can be supported as a sequential equilibrium. In addition, it has similar features to the case without side payments; it is symmetric, and there is no first-mover advantage. Since only low-type regions may delay, an agreement is settled earlier if it is more valuable. And when the regions finally settle the agreement, it is stable, i.e., it coincides with the outcome in (4.2) if information is complete. This sequential equilibrium is, in fact, unique if we restrict the attention to stable outcomes and apply the same *optimistic intuitive criterion* as before.[27] Finally, the sequential equilibrium implements the most efficient stable robust mechanism.[28]

Lemma 2: (i) Equations (4.2)-(4.3) characterize a sequential equilibrium outcome, (ii) which is unique under the *optimistic intuitive criterion* and if we require the outcome to be stable. Moreover, (iii) it implements the most efficient stable robust mechanism.

4.3. When is Uniformity Better?

When both differentiation and side payments are on the negotiation table, the total expected utility can be written as

$$u^{ds} = \frac{1}{4}(\overline{v}-1) + \frac{1}{2}\left[\frac{\overline{v}+\underline{v}}{2} - 1 + (\overline{v}-\underline{v})\left|\frac{1}{2}-e\right| D\right]\delta^{t_1^s} + \frac{1}{4}(\underline{v}-1)\delta^{t_2^s}. \tag{4.4}$$

Gibbons and Klemperer (1987) on how to efficiently solve a partnership. When the parties have roughly equal shares, there is confusion about who is going to sell/buy the shares in the partnership, which makes it easier to encourage a player to reveal its value.

[27] Definitions of these concepts are found in the Appendix.

[28] Once more, we can ask how restrictive the stable- and robust-requirements are. By relaxing the first of these, the best mechanism requires sufficiently large side payments (from the low-type to the high-type region) to make imitation unattractive for the high type. This achieves the first-best. If we only relax the second requirement, the best mechanism dictates no delay if only one region announces low type, but a longer delay if done by both.

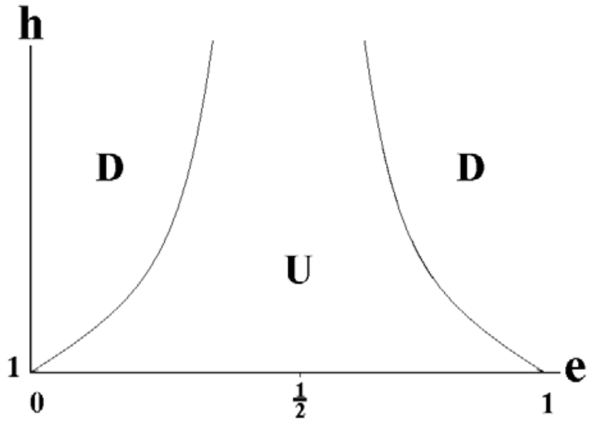

Figure 4.2: *It is optimal to prohibit both differentiation and side payments, instead of allowing them both, if and only if the parameters are such that we are in area U.*

Suppose the two regions were able to commit to uniform policies ($D = 0$) should they ever reach an agreement. Would they make such a commitment? If side payments, but not differentiation, are on the bargaining agenda, the outcome is exactly as above if we set $D = 0$. Define the resulting total expected utility as u^s. By inspection, it is clear that u^{ds} increases in D, for two reasons. First, as D increases, it becomes possible to concentrate more of the cleaning to one region, and the gains from doing this efficiently increases. Second, it becomes more costly for high-type regions to imitate the low-type region's strategy, since this would imply inefficient differentiation. Thus, the need for delay is smaller. For these two reasons, it is always better to allow policy differentiation if side payments are on the agenda.

Proposition 2: $u^{ds} \geq u^s$ always.

This proposition does *not* imply that differentiation is good whenever side payments *can* be part of the agenda. It might be beneficial to prohibit *both* side payments and differentiation, that is, $u^0 \geq u^{ds}$. By doing this, delay is ensured to be zero. By allowing both side payments and differentiation, however, the policy is differentiated efficiently.

It turns out that it is better to prohibit both side payments and differentiation whenever

$$h\left(\frac{2}{1-|1-2e|\,D}-3\right) \leq 3, \tag{4.5}$$

that is, when $|e-1/2|$, h and D are small. The basic intuition is as follows. If $e < 1/2$, in equilibrium, most of the cleaning takes place in the high-value region, and its benefit is decreasing in e. If $e > 1/2$, optimal differentiation implies that the low-type region does most of the cleaning, and this benefit is increasing in e. In either case, the value of such differentiation is increasing in the heterogeneity h and the possible amount of differentiation, D. If $e \approx 1/2$, however, it is of no importance where cleaning takes place, and there is little value of differentiation. The potential cost of differentiation is delay, but this is decreasing in the gains from trade $D\,|1-2e|$, since such trade provides an efficient signaling device. If $D|1-2e| \geq 1$, there is in fact no delay, and the first-best is attainable by allowing both differentiation and side payments.[29]

Proposition 3: $u^0 \geq u^{ds}$ if and only if both $D|1-2e| < 1$ and (4.5) hold. This is more likely if the heterogeneity h is small, the possibilities to differentiate D is small and contributions are almost pure public goods, i.e., $\left|\frac{1}{2}-e\right|$ is small.

5. Are Side Payments Good?

-*Side payments are needed to reach the best result* (Cesar and de Zeeuw, 1996, p.158). By introducing side payments, any outcome raising total welfare can be Pareto improving by making the winner compensate the loser. It is therefore a common presumption that side payments increase the efficiency of negotiations, and economists are eager to advocate issue linkages as a way of introducing side payments.[30] It is therefore puzzling why side payments are seldom observed to be an explicit part of international agreements. Cesar and de Zeeuw conjecture that *the reason might be that it is difficult...to determine the precise willingness to pay.*

Suppose that behind a veil of ignorance, the regions were able to prohibit side payments in future negotiations. Would they? The above analysis makes us well equipped to address this question. The bargaining outcomes for the relevant cases are already discussed: Section 3 made a horizontal comparison between the two agendas in the first row of the table

		Differentiation?	
		no	yes
Side payments?	no	x	d, x
	yes	s, x	d, s, x

Section 4 did a similar comparison between the two agendas in the second row, and compared agenda (d, s, x) to agenda (x). This section exhausts the model by making a vertical comparison.

Introducing side payments to the agenda has three effects. First, it allows one region to compensate the other for contributing more. Such trade is valuable whenever the policy is suboptimally differentiated without side payments. Second, a region can signal its type by the proposed direction of trade. If $e < 1/2$, it is not very attractive for the high-type region to imitate the low-type region's strategy by paying the other region to do most

[29] It should be noticed that the average value v does not influence this condition. Without side payments, a larger v makes a region more willing to contribute (when $e < 1/2$) and less willing to engage in haggling over d. The utility of a high-type region thus increases relative to a low-type region. Side payments adjust to nullify this effect.

[30] E.g. Barrett (2001) writes that side payments can *sustain a vastly superior outcome compared to the agreement without side payments.* From a game-theoretic perspective, however, it is not clear whether side payments are beneficial. Jackson and Wilkie (2003) show that the possibility to commit to side payments conditional on strategies may induce players to inefficiently tilt the equilibrium in their favor. In the present paper, however, agents are not able to make such a commitment prior to the game, and side payments would always be first-best if information were perfect. See also Prat and Rustichini (2003).

of the cleaning. The necessary delay to credibly signal bargaining power is then smaller. There is, however, also a third effect. Without side payments, a low-type region might convince the other region to contribute more. If side payments are allowed, the low-type region may also require transfers from the other region. If so, bargaining power pays off even more and it is more attractive to imitate the low-type's strategy. The *incentives* to signal bargaining power increase, and this effect might outweigh the reduced *necessity* to signal by delay. The resulting delay can either decrease or increase.[31]

Suppose $e < 1/2$. It turns out that side payments are detrimental to total welfare whenever

$$\frac{h-1}{h+1} \geq D\left(\frac{1-v(1-2e)}{v-1}\right), \tag{5.1}$$

that is, when h is large, while D and e are small. The basic intuition is as follows. If D is large, the gains from trade are large, and proposing a certain direction of trade is a credible signal of type. Then, side payments are good as they facilitate trade. If D is low, however, the gains from trade are small, and signaling a certain direction of trade is not a very convincing signal. At the same time, bargaining power is not very useful without side payments, since it is not possible to differentiate the policy to any considerable extent in any case. Introducing side payments, however, allows the low-type region to force the high-type region to pay in side payments what it cannot pay in policy. The incentives to signal bargaining power increase, as does delay. Thus, if D is small, side payments are bad. In other words, unless the existing conflict between the regions is sufficiently large, allowing side payments is detrimental to efficiency since it creates a costly conflict of interest. It follows that excluding side payments is always optimal if the policy must be uniform, i.e. $u^0 > u^s$.

If the heterogeneity h is large, there is a great deal of differentiation d' even without side payments, and the gains from trade $D - d'$ are small. The difference in bargaining power, however, is large, and it is quite likely that the low type will get side payments from the high type. To credibly signal bargaining power, delay must increase when side payments are possible. In this case, efficiency is larger when side payments are prohibited.

If v is large or e is low, regions are more willing to clean domestically, instead of engaging in haggling. A large v makes cleaning more valuable, and a small e makes domestic cleaning more important. Regions are therefore less tempted to signal bargaining power, and delay is reduced. Introducing side payments, however, destroys the peace. Once more, bargaining power becomes valuable. Regions become more attracted to signaling bargaining power in order to tilt the transfers in their direction, and delay increases. Thus, side payments are good only if v is small and e is large. If $e > 1/2$, the policy is suboptimally differentiated without side payments. The gains from trade are then larger, and side payments are more likely to increase efficiency.[32]

[31] This possibility is noticed by observers. Concerning European cooperation, Héritier (2002, p. 186) writes that *If an issue is perceived as redistributive, the decision process rapidly becomes polarized and clear-cut conflict lines emerge... Those adversely affected fend of expected costs and signal their rejection of the proposal.*

[32] It should be noticed that (5.1) is also the condition for when $\underline{s} \leq 0$. Suppose that $d' \geq D$, such that the constraint $d \in [-D, D]$ binds even if side payments are prohibited. Differentiation is then optimal, and there are no gains from trade. Allowing side payments is then beneficial if and only if this reduces

Proposition 4: Suppose that $e \leq 1/2$. $u^d \geq u^{ds}$ if and only if (5.1) holds. This is more likely if the externality e is small, the heterogeneity h large, the value v of the agreement large, while the possibility to differentiate D is small.

6. Interpretations

6.1. International Cooperation

This paper has studied two regions (or countries) trying to coordinate their policies so as to internalize externalities. While policy differentiation is necessary to tie local policies to local conditions, it increases the conflict of interests between the countries, and thus delay when bargaining power is private information. While side payments create gains from trade as well as an efficient signaling device, they may also increase the conflict of interest between the countries, and thus delay. The results described how the best bargaining agenda hinges on the value of an agreement, the externality, the heterogeneity, and the possibilities to differentiate.

The results can be interpreted in several ways. On the positive side, the findings may explain why side payments not always are observed as an explicit part of international agreements,[33] and why federal policies are characterized by uniform policies.[34] On the normative side, the results describe when such harmonization clauses are a good idea, and when it is efficient to allow for issue linkages (e.g. by letting the issue be determined in the European Council, instead of the Council of Ministers that has less discretion). To implement the optimal bargaining agenda, however, a commitment is required in advance, which provides additional interpretations.

6.2. Decentralization vs. Centralization

The above arguments are closely related to the literature on fiscal federalism (surveyed by Oates, 1999). This literature typically compares decentralization vs. centralization of a political instrument under two assumptions. First, the policy is uniform whenever the instrument is centralized. Relying on this assumption, centralization should be a certain way of committing to uniformity. Second, there is no coordination between regions if the policy is decentralized. However, even if the policy is decentralized, the regions have incentives to cooperate whenever externalities exist. According to the Coase Theorem, they will also be quite successful in doing so. If we thus relax the second assumption, the case for decentralization coincides with the case for differentiation analyzed above. If

delay. Delay is reduced if and only if it becomes less tempting for the high type to imitate the low type's strategy. Whether the high type is more or less tempted to imitate the low type depends on whether it will receive or pay side payments. If the high-type region will be compensated for contributing more in equilibrium, then the high type is better off by allowing side payments, and delay is less necessary to make the low type's strategy unattractive for the high type. If, instead, in equilibrium, the high-type region will pay the low-type region, the high-type region is worse off when side payments are allowed, and it becomes more tempting to imitate the low type's strategy. Then, more delay is necessary to credibly signal bargaining power.

[33] This is questioned by e.g. Cesar and de Zeeuw (1996).

[34] Documented by e.g. McCormick (2001).

the policy is decentralized, regions coordinate and differentiate the policy whenever they are heterogeneous. These regional negotiations are likely to be inefficient and delayed, however. Centralizing the policy, instead, gives a clumsy central government no other choice than to implement uniform policies across the regions. While this certainly has a cost when regions differ, the benefit is less delay. Propositions 1 and 3 show that centralization is better if heterogeneity is low and the externality large. While this is in line with the traditional literature, the results also provide new recommendations. Proposition 1 shows that differentiation is better when the value of an agreement is large, because delay is then smaller. Hence, more important decisions should be decentralized. Moreover, in contrast to the earlier literature, I find that it is the existence of asymmetric information which makes the case for centralization. With complete information, it is always better to differentiate the policy, and thus allow decentralized coordination. With asymmetric information, instead, decentralized coordination is likely to be inefficient and centralization may be better. Finally, the central government's uniform policies do not constitute a disadvantage, calling for more decentralization (as normally argued). In contrast, it is the uniform policy which makes centralization potentially attractive, since it reduces the transaction costs.

The analysis also suggests a case for partial decentralization. Comparing Propositions 1 and 4, a differentiated policy might be better than a uniform one, but side payments may still be a bad idea. This will typically be the case if heterogeneity is large while there is a limit to how much it is possible to differentiate the policy. The best political regime is then to decentralize the relevant policy while restricting the regions' discretion over side payments.

6.3. Integration and Uniformity

As described by Propositions 1-4, regions may benefit if negotiating uniform instead of differentiated policies. As noticed, however, this requires commitment. Without commitment, a reluctant region can easily propose differentiation and perhaps side payments. One way of committing is to use trigger strategies in frequent interaction, where regions stick to the restricted bargaining agenda (without differentiation or side payments) if this facilitates future cooperation. Another way of committing is to write formal agreements, calling for harmonized policies. For either reason, regions constituting a federal union should be better able to commit to a restricted agenda when this is the best solution. Hence, we should observe more uniform policies between regions forming a federal union than between regions that do not. This is exactly the first assumption, mentioned above, made by the traditional literature on fiscal federalism. The above analysis thus provides a theoretical foundation for this, and characterizes when the uniformity assumption is likely to hold. According to this argument, however, this uniformity is not a necessary shortcoming due to the central government's inability to differentiate, as claimed by the fiscal federalism literature. In contrast, the uniformity is a benefit arising as the federal union makes the regions better able to commit.

6.4. Status Quo Bias

Quite often, constitutional rules make renegotiation costly. A justification for such status-quo bias is provided by the analysis of this paper. In fact, a commitment to uniform policies may be interpreted as a commitment to stick to an agreement settled behind a veil of ignorance. One certain way of committing is to create obstacles to renegotiations. While the cost of this is that the policy cannot be optimally differentiated ex post, once the types are realized, the benefit is that regions will not undertake distorting signaling and screening to tilt the agreement to their advantage. Even if regions were allowed to renegotiate the allocation of contributions, a status quo bias on other political issues may effectively prevent the introduction of side payments. Thus, Propositions 1-4 above can alternatively be interpreted as conditions making a status quo bias rational.

6.5. Further Research

Designing a constitution permits more than banning differentiation and side payments. More importantly, a constitution defines how future decisions will be taken, that is, the rules of the bargaining game. Investigating how different constitutional rules are able to mitigate the inefficiencies described in this paper is an interesting issue for future research. This raises a host of questions. To which extent, for example, is it a good idea to concentrate the agenda-setting power to one region? How does the optimal constitution change when the number of regions increases? What is the optimal majority rule?

The general lesson of this paper is that parties negotiating under private information may benefit from simply constraining the agenda. With a great deal of discretion, a strong party is fully able to exploit its bargaining power. It is then very beneficial to signal bargaining power and screen the other party. Typically, this creates distortions. By instead restricting the agenda, the conflict of interest between the parties may decrease, it becomes more difficult to exploit bargaining power, and distortions diminish. If the value of discretion is small, efficiency benefits from constraining the agenda.

This trade-off between flexibility and costly signaling can be applied to many contexts. Take the Theory of the Firm. By definition, a market transaction requires a price and thus, a conflict between the seller and the buyer. It is no surprise that most bargaining theory is developed for such situations. If the transaction were undertaken in-house, however, the incentives might be less conflicting. In fact, it can be argued that in-house transactions are insensitive to whether the realized benefit is larger than the realized cost. These pieces of information may be private to different employees with small incentives to coordinate. But, as they are not haggling, delay is reduced. The traditional theory of the firm, such it is surveyed by Hart (1995), emphasizes how ownership affects incentives prior to negotiations. It might be time to turn the attention to ex post transaction costs. When is it good to forbid side payments within firms? Which transactions are better undertaken within instead of between firms?

Appendix

PROOF OF LEMMA 1

Proof of (i) and (ii): At any point in time, a *history* after N offers is the set of proposed and rejected offers: $H_N = \{d_N, t_N\}_N$. Let $\mathbf{H}_N$ denote the set of such possible histories, define $H_0 \equiv (0,0)$, and let $\mathbf{H}$ be the set of all possible histories (any N). A *pure strategy* for A is a rule f_A that says, whenever N is even, whether A should accept the previous offer or make a counteroffer d_{N+1} after some delay $t_{N+1} - t_N \geq 0$; that is, $f_A : \mathbf{H} \longrightarrow \{accept, (\mathbb{R}, \mathbb{R}_+)\}$. Let A's *belief* $b_A : \mathbf{H} \longrightarrow [0,1]$ denote the probability A puts on the event $v_B = \underline{v}$ after some history H_N. Similarly, f_B and b_B denote B's strategy and beliefs about A's type. At time $t = 0$, $b_A = b_B = 1/2$.

A *sequential equilibrium* (Kreps and Wilson, 1982) is a set of strategies and beliefs such that after every history, each player's strategy is optimal, given its beliefs and the other player's strategy, and the beliefs are consistent with Bayes' rule. The *intuitive criterion* (Cho and Kreps 1987) is a refinement which puts restrictions on beliefs outside the equilibrium. In essence, it requires that any action out of equilibrium beneficial for exactly one type, implies that beliefs place probability one on this type. To ensure a unique equilibrium in the above game in a simple way, I will apply an even stronger updating rule.

Definition 4: Let $(d,t)_i$ denote an (expected) outcome if i is of high type, given i's belief. Let $F_i \equiv \{(d,t) | (d,t) \succ_i (d,t)_i$ if and only if $v_i = \underline{v}\}$. The *intuitive criterion* requires that $b_j = 1$ after $i \neq j$ has taken some action leading to an outcome in F_i. In addition, the *optimistic intuitive criterion* requires that $b_i = 0$ if i has taken some action leading to an outcome outside F_i.

This criterion requires that after a region has made an offer, unless this offer is unattractive for the high-type region, the region is believed (by the other region) to be of high type for certain. This way of updating beliefs is quite "optimistic", though certainly possible.

Suppose A is revealed to be of low type by making an offer at $t_{\underline{A}}$. A high-type B will not be able to convince A that B is of low type. Thus, B accepts any $d \leq d'$, and will itself immediately propose d' if A's proposal is some $d > d'$ (remember that d' is the equilibrium when $b_A = 0$ and $b_B = 1$ are correct beliefs). A low-type B, on the other hand, maximizes its utility by proposing an offer in F_B which is acceptable by A if $b_A = 1$; that is, it must be unattractive to a high-type B and acceptable to the low-type A with beliefs $b_A = 1$:

$$\underset{(d,t_2)}{Max} \; \frac{1}{2}\left[\underline{v} - 1 - (1 - \underline{v}(1-2e))\, d\right] \delta^{t_2} \quad \text{s.t.}$$

$$\begin{aligned} \frac{1}{2}\left[\overline{v} - 1 - (1 - \overline{v}(1-2e))\, d'\right] \delta^{t_{\underline{A}}} &\geq \frac{1}{2}\left[\overline{v} - 1 - (1 - \overline{v}(1-2e))\, d\right] \delta^{t_2} \\ d &\geq 0. \end{aligned}$$

The solution is

$$\begin{aligned} d &= 0 \\ \delta^{t_2 - t_{\underline{A}}} &= \frac{\overline{v} - 1 - (1 - \overline{v}(1-2e))\, d'}{\overline{v} - 1}. \end{aligned}$$

Suppose instead that A is revealed to be of high type by making an offer at $t_{\overline{A}}$. A high-type B will not be able to convince A that B is of low type, and B accepts any

$d \leq 0$, and will itself immediately propose $d = 0$ if A's proposal is some $d > 0$ (remember that $d = 0$ is the equilibrium when $b_A = 0$ and $b_B = 0$ are correct beliefs). A low-type B, on the other hand, maximizes its utility by proposing an acceptable offer in F_B; that is, it must be unattractive to a high-type B and acceptable to a high-type A with beliefs $b_A = 1$:

$$\underset{(d,t_{\underline{B}})}{Max} \ \frac{1}{2}\left[\underline{v} - 1 - (1 - \underline{v}(1-2e))\, d\right] \delta^{t_{\underline{B}}} \ \text{ s.t.}$$

$$\begin{aligned} \frac{1}{2}(\overline{v}-1)\,\delta^{t_{\overline{A}}} &\geq \frac{1}{2}\left[\overline{v} - 1 - (1-\overline{v}(1-2e))\, d\right] \delta^{t_{\underline{B}}} \\ d &\geq -d'. \end{aligned}$$

The solution is

$$\begin{aligned} d &= -d' \\ \delta^{t_{\underline{B}} - t_{\overline{A}}} &= \frac{\overline{v}-1}{\overline{v} - 1 + (1-\overline{v}(1-2e))\, d'}. \end{aligned}$$

Having found B's optimal strategy, let us turn to A. If A is of high type, it can (by Definition 4) not afford to persuade B to believe that $b_B \neq 0$. Thus, A can either make a pooling offer $-d'$ which is acceptable to B whatever its type, or A can make a screening offer $d = 0$ which will only be accepted by a high-type B. Since we know B's reaction in either case, it is easily calculated that A is better off by making the screening offer $d = 0$ at $t_{\overline{A}} = 0$. This gives A the expected utility

$$\overline{u}_A = \frac{1}{4}(\overline{v}-1) + \frac{1}{4}\left[\overline{v} - 1 - (1-\overline{v}(1-2e))\, d'\right] \delta^{t_{\underline{B}}}.$$

The low-type A's problem is then to make an offer which is not attractive to a high-type A, but acceptable to a high-type B with beliefs $b_B = 1$ (it can easily be shown that a low-type A will not make a pooling offer):

$$\underset{(d,t_{\underline{A}})}{Max} \ \frac{1}{4}\left[\underline{v} - 1 + (1-\underline{v}(1-2e))\, d\right] \delta^{t_{\underline{A}}} + \frac{1}{4}(\underline{v}-1)\,\delta^{t_2 - t_{\underline{A}}}\delta^{t_{\underline{A}}} \ \text{ s.t.}$$

$$\begin{aligned} \overline{u}_A &\geq \frac{1}{4}\left[\overline{v} - 1 + (1-\overline{v}(1-2e))\, d\right] \delta^{t_{\underline{A}}} + \frac{1}{4}(\overline{v}-1)\,\delta^{t_2 - t_{\underline{A}}}\delta^{t_{\underline{A}}} \\ d &\leq d'. \end{aligned}$$

The solution can be shown to be

$$\begin{aligned} d &= d' \\ \delta^{t_{\underline{A}}} &= \frac{\overline{v}-1}{\overline{v} - 1 + (1-\overline{v}(1-2e))\, d'}. \end{aligned}$$

Combined, it follows that $\delta^{t_1} \equiv \delta^{t_{\underline{A}}} = \delta^{t_{\underline{B}}}$ and δ^{t_2} are such as these are defined in (3.4). In equilibrium, the low type will delay and the high type will not. The beliefs in the optimistic intuitive criterion are therefore consistent with Bayes' rule. It follows that these beliefs and the strategies above constitute a sequential equilibrium, which is unique under the optimistic intuitive criterion. Since strategies are symmetric in the sense that $\delta^{t_{\underline{A}}} = \delta^{t_{\underline{B}}}$, and since B acts after A's type has been revealed, the low-type A's strategy is not attractive to the high-type A, even if A had "spied" on B and had beliefs $b_A \in \{0, 1\}$. Thus, the mechanism is transparent. It can also be shown that no pooling equilibrium exists under a weak form of the intuitive criterion (where $b_j = 1/2$, unless i's action is in F_i).[35]

Proof of (iii): According to Definition 2, a *mechanism* is a rule mapping any pair of announced types to an outcome (d, t). Readers preferring to design static mechanism in terms of the probability of agreement, instead of time, can simply let δ^{t_i} denote this probability. I will now calculate the most efficient mechanism that is stable ($d = d_R$) and incentive compatible, even if a region should be aware of the other region's announcement (i.e. robust). This mechanism maximizes the total expected utility by minimizing delay, subject to these constraints and the regions' incentive constraints. The participation constraints are fulfilled when $d = d_R$. Let t_0, t_1 and t_2 denote the time of the settlement when, respectively, none, one and both regions announce low type. Since the game is symmetric, I do not need to let t_1 depend on which of the regions announces low type (this would not change the result). The problem is

$$\underset{t_0,t_1,t_2\in[0,\infty)}{Max} u^d = \frac{1}{4}(\overline{v}-1)\delta^{t_0} + \frac{1}{4}(\underline{v}-1)\delta^{t_2} + \frac{1}{2}\left[\frac{1}{2}(\underline{v}+\overline{v}) - 1 + \left(\frac{1}{2}-e\right)(\overline{v}-\underline{v})d'\right]\delta^{t_1} \quad s.t.$$

$$\frac{1}{2}(\overline{v}-1)\delta^{t_0} \geq \frac{1}{2}[\overline{v}-1+(1-\overline{v}(1-2e))d']\delta^{t_1} \qquad (\overline{IC})$$

$$\frac{1}{2}[\overline{v}-1-(1-\overline{v}(1-2e))d']\delta^{t_1} \geq \frac{1}{2}(\overline{v}-1)\delta^{t_2}, \qquad (\underline{IC})$$

where $(\overline{IC})$ and $(\underline{IC})$ are the high type's incentive constraints when the other region announces high and low type, respectively. When $(\overline{IC})$ and $(\underline{IC})$ both hold, truthful announcement becomes optimal also if a region is uncertain about the other region's type. It is easily checked that the low type's incentive constraints are not binding, and these can therefore be ignored. The solution is that $t_0 = 0$, while t_1 and t_2 are set such that

$$\delta^{t_1} = \frac{\overline{v}-1}{\overline{v}-1+(1-\overline{v}(1-2e))d'}$$

$$\delta^{t_2} = \frac{\overline{v}-1-(1-\overline{v}(1-2e))d'}{\overline{v}-1+(1-\overline{v}(1-2e))d'} \Leftrightarrow (3.4).$$

PROOF OF PROPOSITION 1: Define the *net* values of a uniform agreement as $\underline{n} \equiv \underline{v} - 1$ and $\overline{n} \equiv \overline{v} - 1$. Note that $\overline{n} = 2(v-1)h/(h+1)$ and $\left(1-\delta^{t_2}\right) = 2\left(1-\delta^{t_1}\right)$.

[35] In a previous version of this paper, the probability that $v_i = \underline{v}$ could be $p > 1/2$. Then, pooling offers might be optimal and pooling equilibria where all types suggest a uniform policy $d = 0$ might exist. Uniform policies might then be the outcome even if heterogeneous regions are allowed to differentiate the policy. However, the set of parameters under which such pooling equilibria exist is strictly smaller than the set of parameters under which a commitment to uniform policies would be good. The results of the paper thus survive, while the analysis would be more complicated.

By comparing (3.5) and (3.6):

$$\begin{aligned}
u^d &\leq u^0 \Leftrightarrow \\
\frac{1}{2}\left(\frac{1}{2}-e\right)(\overline{v}-\underline{v})d'\delta^{t_1} &\leq \frac{1}{2}\left(\frac{\underline{v}+\overline{v}}{2}-1\right)\left(1-\delta^{t_1}\right)+\frac{1}{4}(\underline{v}-1)\left(1-\delta^{t_2}\right) \Leftrightarrow \\
(1-2e)(\overline{v}-\underline{v})d' &\leq (\underline{v}+\overline{v}-2)\left(1-\delta^{t_1}\right)/\delta^{t_1}+(\underline{v}-1)2\left(1-\delta^{t_1}\right)/\delta^{t_1} \Leftrightarrow
\end{aligned}$$

$$\begin{aligned}
(1-2e)(\overline{v}-\underline{v})d' &\leq (3\underline{v}+\overline{v}-4)\frac{[1-\overline{v}(1-2e)]d'}{\overline{v}-1} \Leftrightarrow \\
(1-2e)(\overline{n}-\underline{n})\overline{n} &\leq (3\underline{n}+\overline{n})[1-(\overline{n}+1)(1-2e)] \Leftrightarrow
\end{aligned}$$

$$\begin{aligned}
(1-2e)(2\overline{n}+2\underline{n})\overline{n} &\leq (3\underline{n}+\overline{n})2e \Leftrightarrow \\
(1-2e)2(v-1)h &\leq (3+h)e \Leftrightarrow \\
h[2(1-2e)(v-1)-e] &\leq 3e \Leftrightarrow (3.7).
\end{aligned}$$

PROOF OF LEMMA 2

Proof of (ii): This proof is similar to the proof of Lemma 1. With side payments, a *history* after N offers is the set of proposed and rejected offers: $H_N = \{d_N, s_N, t_N\}_N$. A *pure strategy* for A is a rule $f_A : \mathbf{H} \longrightarrow \{accept, ([-D, D], \mathbb{R}, \mathbb{R}^+)\}$. Let $(d,s,t)_i$ denote an (expected) outcome if i is of high type, given i's belief. Let $F_i \equiv \{(d,s,t)|(d,s,t) \succ_i (d,s,t)_i$ if and only if v_i The *intuitive criterion* requires that $b_j = 1$ after $i \neq j$ has taken some action leading to an outcome in F_i. In addition, the *optimistic intuitive criterion* requires that $b_i = 0$**,** if i has taken some action leading to an outcome outside F_i. The attention will be restricted to stable offers (Definition 1).[36]

Suppose $e \leq 1/2$, and that A is revealed to be of low type by making an offer at $t^s_{\underline{A}}$. A high-type B will not be able to convince A that B is of low type, and will propose $d = D$ and $s = \underline{s}$, giving B utility $\overline{u}_B = [\overline{v}+\underline{v}-2+(\overline{v}-\underline{v})(1-2e)D]/4$. In considering A's offer, a high-type B accepts anything that would make B's utility at least as large as $\overline{u}_B$. A low-type B, on the other hand, prefers to propose a stable agreement $(0,0,t^s_2) \in F_B$ which is thus acceptable to A

$$\underset{t^s_2 \geq t^s_{\underline{A}}}{Max} \ \frac{1}{2}(\underline{v}-1)\delta^{t^s_2} \text{ s.t.}$$

$$\overline{u}_B \delta^{t^s_{\underline{A}}} \geq \frac{1}{2}(\overline{v}-1)\delta^{t^s_2}.$$

The solution is

$$\begin{aligned}
\delta^{t^s_2 - t^s_{\underline{A}}} &= \frac{\overline{v}+\underline{v}-2+(\overline{v}-\underline{v})(1-2e)D}{2(\overline{v}-1)} \text{ if } (1-2e)D < 1 \\
t^s_2 &= t^s_{\underline{A}} \text{ if } (1-2e)D \geq 1.
\end{aligned}$$

[36] Wang (2000) also restricts the attention to stable offers in his derivation of unique equilibria in the Cramton (1992) model. See the next footnote for my justification for restricting the attention to stable offers.

Suppose instead that A is revealed to be of high type by making an offer at $t^s_{\overline{A}}$. A high-type B will not be able to convince A that B is of low type, and will propose $d = s = 0$, giving B utility $(\overline{v} - 1)/2$. In considering A's offer, a high-type B accepts anything that would make B's utility at least as large as $(\overline{v} - 1)/2$. A low-type B, on the other hand, prefers to propose the stable agreement $(-D, -\underline{s}, t^s_B) \in F_B$ which is thus acceptable to A.

$$\underset{t^s_B \geq t^s_{\overline{A}}}{Max} \; \frac{1}{2}\left[\underline{v} - 1 + (1 - \underline{v}(1-2e))\, D + \underline{s}\right] \delta^{t^s_B} \quad \text{s.t.}$$

$$\frac{1}{2}(\overline{v} - 1)\,\delta^{t^s_{\overline{A}}} \geq \frac{1}{2}\left[\overline{v} - 1 + (1 - \overline{v}(1-2e))\, D + \underline{s}\right] \delta^{t^s_B}.$$

Substituting for $\underline{s}$, the solution becomes

$$\begin{aligned} \delta^{t^s_B - t^s_{\overline{A}}} &= \frac{\overline{v} - 1}{\overline{v} - 1 + (\overline{v} - \underline{v})(1 - (1-2e)D)/2} \quad \text{if } (1-2e)\, D < 1 \\ t^s_B &= t^s_{\overline{A}} \quad \text{if } (1-2e)\, D \geq 1. \end{aligned}$$

Having found B's optimal strategy, let us turn to A. If A is of high type, it cannot afford to persuade B to believe that $b_B \neq 0$. Thus, A can either make a pooling offer which is acceptable to B whatever its type, or A can make a screening offer which only a high-type B would accept. Since we know B's reaction in either case, it is easy to calculate that A is better off by making the screening offer $d = s = 0$ at $t^s_{\overline{A}} = 0$. This gives A the expected utility

$$\overline{u}_A = \frac{1}{4}(\overline{v} - 1) + \frac{1}{4}\left[\overline{v} - 1 - (1 - \overline{v}(1-2e))\, D - \underline{s}\right] \delta^{t^s_B}.$$

The low-type A's problem is then to make an offer which is not attractive to a high-type A, but acceptable to a high-type B with beliefs $b_B = 1$. It can easily be shown that a low-type A will not make a pooling offer, so A proposes the stable agreement $\left(D, \underline{s}, t^s_{\underline{A}}\right) \in F_A$

$$\underset{t^s_{\underline{A}} \geq 0}{Max} \; \frac{1}{4}\left[\underline{v} - 1 + (1 - \underline{v}(1-2e))\, D + \underline{s}\right] \delta^{t^s_{\underline{A}}} + \frac{1}{4}(\underline{v} - 1)\,\delta^{t^s_2 - t^s_{\underline{A}}}\delta^{t^s_{\underline{A}}} \quad \text{s.t.}$$

$$\overline{u}_A \geq \frac{1}{4}\left[\overline{v} - 1 + (1 - \overline{v}(1-2e))\, D + \underline{s}\right] \delta^{t^s_{\underline{A}}} + \frac{1}{4}(\overline{v} - 1)\,\delta^{t^s_2 - t^s_{\underline{A}}}\delta^{t^s_{\underline{A}}}.$$

The solution can be shown to be

$$\begin{aligned} \delta^{t^s_{\underline{A}}} &= \frac{\overline{v} - 1}{\overline{v} - 1 + (\overline{v} - \underline{v})(1 - (1-2e)D)/2} \quad \text{if } (1-2e)\, D < 1 \\ t^s_{\underline{A}} &= 0 \quad \text{if } (1-2e)\, D \geq 1. \end{aligned}$$

If $e > 1/2$, the proof proceeds in the same way, but since d changes signs in the optimal agreement, $(1-2e)$ should be replaced by $|1-2e|$. Combined, it follows that $\delta^{t^s_1} \equiv \delta^{t^s_{\underline{A}}} = \delta^{t^s_B}$ and $\delta^{t^s_2}$ are such as these are defined in (4.3). In equilibrium, the low type will delay and the high type will not. The beliefs under the optimistic intuitive criterion are therefore consistent with Bayes' rule. It follows that these beliefs and the above strategies comprise a unique equilibrium under the optimistic intuitive criterion when the agreement must be stable.[37] It can also be shown that no pooling equilibrium exists under a weak form of the intuitive criterion (where $b_j = 1/2$ unless i's action is in F_i).[38]

Proof of (i): The above equilibrium is shown to be unique under the optimistic intuitive criterion and when the agreement is stable. Thus, it is a sequential equilibrium if proposals must be stable. If beliefs are such that $b_i = 0$ whenever $j \neq i$ proposes an agreement which is not stable, it is easily shown that the above strategies constitute a sequential equilibrium even if offers do not have to be stable.

Proof of (iii): I will now calculate the most efficient mechanism that is stable and robust. This mechanism maximizes the total expected utility by minimizing delay subject to these constraints and the regions' incentive constraints. The participation constraints are fulfilled when the agreement is stable. Let t_0^s, t_1^s and t_2^s denote the time of the settlement when none, one and both regions announce low type, respectively. Since the game is symmetric, I do not need to let t_1^s depend on which of the regions announces low type (doing this would not change the result). Suppose $e \geq 1/2$. The problem is

$$\underset{t_0^s, t_1^s, t_2^s \geq 0}{Max} \quad u^{ds} = \frac{1}{4}(\overline{v} - 1)\delta^{t_0^s} + \frac{1}{2}\left[\frac{\underline{v} + \overline{v}}{2} - 1 + (\overline{v} - \underline{v})\left|\frac{1}{2} - e\right| D\right]\delta^{t_1^s} + \frac{1}{4}(\underline{v} - 1)\delta^{t_2^s} \text{ s.t.}$$

$$\frac{1}{2}(\overline{v} - 1)\delta^{t_0^s} \geq \frac{1}{2}\left[\overline{v} - 1 + (1 - \overline{v}(1 - 2e))D + \underline{s}\right]\delta^{t_1^s} \quad (\overline{IC})$$

$$\frac{1}{2}\left[\overline{v} - 1 - (1 - \overline{v}(1 - 2e))D - \underline{s}\right]\delta^{t_1^s} \geq \frac{1}{2}(\overline{v} - 1)\delta^{t_2^s} \quad (\underline{IC})$$

where $(\overline{IC})$ and $(\underline{IC})$ are the high type's incentive constraints when the other region is of high and low type, respectively. When $(\overline{IC})$ and $(\underline{IC})$ both hold, truthful announcements are optimal also if a region is uncertain about the other region's type. It is easily checked that the low type's incentive constraints are not binding. Substituting for $\underline{s}$, it follows that $t_0^s = 0$, while t_1^s and t_2^s are set such that:

$$\begin{aligned}
\delta^{t_1^s} &= \frac{\overline{v} - 1}{(\overline{v} - 1) + \frac{1}{2}(\overline{v} - \underline{v})(1 - (1 - 2e)D)} \quad \text{and} \\
\delta^{t_2^s} &= \frac{(\overline{v} - 1) - \frac{1}{2}(\overline{v} - \underline{v})(1 - (1 - 2e)D)}{(\overline{v} - 1) + \frac{1}{2}(\overline{v} - \underline{v})(1 - (1 - 2e)D)} \quad \text{if } (1 - 2e)D < 1. \\
t_0^s &= t_1^s = t_2^s = 0 \quad \text{if } (1 - 2e)D \geq 1.
\end{aligned}$$

If $e > 1/2$, a similar maximization problem gives the same solution if only $(1 - 2e)$ is replaced by $|1 - 2e|$.

[37] Why require the agreement to be stable? If A is proved to be of low type, a low-type B could save delay by proposing $d = -D$ instead of $d = 0$, by adjusting the side payments accordingly (to equalize utilities). If small transaction costs were related to the side payments, however, A and B would prefer to renegotiate and set $d = s = 0$, when both are proved to be of low type. Hence, signaling by proposing $d = -D$ would not be credible, since the agreement would not be stable.

[38] In a previous version of this paper, the probability that $v_i = \underline{v}$ could be $p > 1/2$. Then, pooling offers might be optimal, and pooling equilibria where all types suggest $d = s = 0$ might exist. Uniform policies might then be the outcome even if heterogeneous regions are allowed to differentiate the policy and negotiate over side payments. However, the set of parameters under which such pooling equilibria exist is strictly smaller than the set of parameters under which a commitment would be good. The results of the paper thus survive, while the analysis would be more complicated.

PROOF OF PROPOSITION 3: If $D\,|1-2e| \geq 1$, we know that the policy is optimally differentiated with no delay. Therefore, assume that $D\,|1-2e| < 1$, and apply the same definitions of $\underline{n}$ and $\overline{n}$ as in the proof of Proposition 1. $u^0 \geq u^{ds}$ whenever the benefit from an optimally differentiated policy is smaller that the cost of delay:

$$\begin{aligned}
\frac{1}{2}\left((\overline{v}-\underline{v})\left|\frac{1}{2}-e\right| D\right)\delta^{t_1^s} &\leq \frac{1}{2}\left(\frac{\underline{v}+\overline{v}}{2}-1\right)(1-\delta^{t_1^s})+\frac{1}{4}(\underline{v}-1)(1-\delta^{t_2^s}) \Leftrightarrow \\
(\overline{v}-\underline{v})\,|1-2e|\, D &\leq (\underline{v}+\overline{v}-2)\frac{1-\delta^{t_1^s}}{\delta^{t_1^s}}+(\underline{v}-1)\frac{2(1-\delta^{t_1^s})}{\delta^{t_1^s}} \Leftrightarrow \\
(\overline{v}-\underline{v})\,|1-2e|\, D &\leq [3\underline{v}+\overline{v}-4]\,\frac{(\overline{v}-\underline{v})\,(1-|1-2e|\, D)}{2\,(\overline{v}-1)} \Leftrightarrow \\
2\overline{n}\,|1-2e|\, D &\leq [3\underline{n}+\overline{n}]\,(1-|1-2e|\, D) \Leftrightarrow \\
\overline{n}(3\,|1-2e|\, D-1) &\leq 3\underline{n}\,(1-|1-2e|\, D) \Leftrightarrow \\
h\left(\frac{3\,|1-2e|\, D-1}{1-|1-2e|\, D}\right) &\leq 3 \Leftrightarrow (4.5).
\end{aligned}$$

PROOF OF PROPOSITION 4: Side payments are beneficial if and only if $u^{sd} \geq u^d$, which requires:

$$\begin{gathered}
[\underline{v}+\overline{v}-2+(\overline{v}-\underline{v})\,|1-2e|\, D]\,\delta^{t_1^s}+(\underline{v}-1)\delta^{t_2^s} \geq \\
[\underline{v}+\overline{v}-2+(\overline{v}-\underline{v})\,(1-2e)\, d']\,\delta^{t_1}+(\underline{v}-1)\delta^{t_2}.
\end{gathered}$$

By introducing side payments, there are always gains from trade since $|1-2e|\, D \geq (1-2e)\, d'$. Suppose that $e \leq 1/2$ and $d' < D$. Comparing (3.4) and (4.3), we notice that side payments reduce delay whenever $(\overline{v}-\underline{v})\,(1-|1-2e|\, D)\,/2 < [1-\overline{v}\,(1-2e)]\, d'$. Substituting for d', we observe that this condition always holds! Suppose, therefore, that $d' \notin [-D, D]$. When this condition binds, d' should be substituted in equilibrium by D. Then, there are no gains from trade, and side payments are good if and only if they reduce delay. The condition for this is:

$$\begin{aligned}
(\overline{v}-\underline{v})\,(1-(1-2e)\, D) &\leq 2\,[1-\overline{v}\,(1-2e)]\, D \Leftrightarrow \\
(\overline{v}-\underline{v}) &\leq D\,[2-(\overline{v}+\underline{v})\,(1-2e)] \Leftrightarrow
\end{aligned}$$

$$h-1 \leq D\left[\frac{2e(h+1)}{v-1}-(h+1)\,(1-2e)\right] \Leftrightarrow (5.1).$$

This condition will always be satisfied when $d' < D$. If $e > 1/2$, the requirement for when side payments reduce delay is relaxed. In addition, the gains from trade are larger. Thus, the larger is e, the more likely are side payments to increase efficiency.

References

Admati, A. and M. Perry (1987): "Strategic Delay in Bargaining", *Review of Economic Studies* 54: 345-364.

Alesina, A. and A. Drazen (1991): "Why are Stabilizations Delayed?", *American Economic Review* 81 (5): 1170-1188.

Alesina, A. and E. Spoalore (1997): "On the Number and Size of Nations", *Quarterly Journal of Economics* 112: 1027-1056.

Ausubel, L., P. Cramton, and R. Deneckere (2002): "Bargaining with Incomplete Information", in R. J. Aumann and S. Hart (Eds.), *Handbook of Game Theory*, Vol. 3, Amsterdam: Elsevier Science B.V., Chapter 50: 1897-1945.

Barrett, S. (2001): "International Cooperation for Sale", *European Economic Review* 45: 1835-1850.

Besharov, G. (2002): "Influence Costs in the Provision of Local Public Goods", Mimeo, Duke University.

Besley, T. and S. Coate (2003): "Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Analysis", *Journal of Public Economics* 87 (12): 2611-37.

Bolton, P. and J. Farrell (1990): "Decentralization, Duplication, and Delay", *Journal of Political Economy* 98 (4): 803-26.

Bolton, P. and G. Roland (1997): "The Break-up of Nations: A Political Economy Analysis", *Quarterly Journal of Economics* 112 (4): 1057-90.

Bordignon, M., P. Manasse and G. Tabellini (2001): "Optimal Regional Redistribution Under Asymmetric Information", *American Economic Review* 91 (3): 709-23.

Cesar, H., and A. de Zeeuw (1996): "Issue Linkage in Global Environmental Problems", in A. Xepapadeas (Ed.), *Economic Policy for the Environment and Natural Resources: Techniques for the Management and Control of Pollution*, Cheltenham, E. Elgar.

Cho, I. and D. Kreps (1987): "Signaling Games and Stable Equilibria", *Quarterly Journal of Economics* 102: 179-221.

Cramton, P. (1992): "Strategic Delay in Bargaining with Two-Sided Uncertainty", *Review of Economic Studies* 59: 205-25.

Cramton, P., R. Gibbons and P. Klemperer (1987): "Dissolving a Partnership Efficiently", *Econometrica* 55: 615-32.

Cremer, J. and T. Palfrey (2000): "Federal Mandates by Popular Demand", *Journal of Political Economy* 108 (5): 905-27.

Ellingsen, T. (1998): "Externalities vs Internalities: A Model of Political Integration", *Journal of Public Economics* 68: 251-68.

Hart, O. (1995): *Firms, Contracts, and Financial Structure*, Oxford University Press.

Héritier, A. (2002): "The Accommodation of Diversity in European Policy-making and its Outcomes: Regulatory Policy as a Patchwork", in A. Jordan (Ed.), *Environmental Policy in the European Union. Actors, Institutions and Processes*, Earthscan Publications Ltd.

Hoel, M. (1992): "International Environment Conventions: The Case of Uniform Reductions of Emissions", *Environmental and Resource Economics* 2: 141-59.

Hsieh, C. (2000): "Bargaining over Reform", *European Economic Review* 44: 1659-76.

Inderst, R. (2000): "Multi-issue Bargaining with Endogenous Agenda", *Games and Economic Behavior* 30: 64-82.

Jackson, M. and S. Wilkie (2003): "Endogenous Games and Mechanisms: Side Payments Among Players", *Review of Economic Studies*, forthcoming.

Kreps, D. and R. Wilson (1982): "Sequential Equilibria", *Econometrica* 50 (4): 863-94.

Lockwood, B. (2002): "Distributive Politics and the Benefits of Decentralisation", *Review of Economic Studies* 69: 313-38.

McCormick, J. (2001): *Environmental Policy in the European Union*, Palgrave.

Myerson, R. and M. Sattertwhite (1983): "Efficient Mechanisms for Bilateral Trading", *Journal of Economic Theory* 29 (2): 265-81.

Mäler, K. G. (1991): "Environmental Issues in the New Europe", in Atkinson, A. B., and R. Brunetta (Eds.), *Economics for the New Europe*, New York University Press.

Oates, W. (1972): *Fiscal Federalism*, Harcourt Brace: New York.

Oates, W. (1999): "An Essay on Fiscal Federalism", *Journal of Economic Literature* 37: 1120-49.

Panagariya, A. and D. Rodrik (1993): "Political-Economy Arguments for a Uniform Tariff", *International Economic Review* 34 (3): 685-703.

Persson, T. and G. Tabellini (1995): "Double-Edged Incentives: Institutions and Policy Coordination", *Handbook of International Economics*, Vol. III.

Prat, A. and A. Rustichini (2003): "Games Played Through Agents", *Econometrica* 71 (4): 989-1026.

Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model", *Econometrica* 50 (1): 97-110.

Wang, R. (2000): "Separating Equilibria in a Continuous-time Bargaining Model with Two-Sided Uncertainty", *International Journal of Game Theory* 29: 229-40.