On delegation, legitimacy, and state building: a signaling approach

van der Weele, J.J.

DOI:
10.2139/ssrn.1724748

On delegation, legitimacy, and state building. A signaling approach.

December 9, 2010

1 Introduction

Governance and state building in post-war societies occupied by foreign militaries or peace-keeping troops is a topic of great practical and theoretical interest. Recent history has born out the monumental difficulties associated with rebuilding a country's institutions when existing structures have been destroyed. A central difficulty facing state builders is to establish the 'legitimacy' of transitional governance structures (Paris and Sisk 2009). Legitimacy refers to the belief by the population that a ruler or institution ought to be obeyed (Hurd 1999: 381) and is associated with trust and cooperation from the local population. The RAND Corporation (2007)'s "Beginner's guide to nation-building" describes how crucial legitimacy is to the success of state building.

> "[T]he symptoms of failed legitimacy are well known: sagging support in public opinion surveys, rampant corruption, economic stagnation, unsuccessful foreign policy initiatives, and loss of support from domestic constituencies such as religious authorities."
>
> RAND Corporation, "Beginner's guide to nation-building" (2007: 194).

Establishing legitimacy is not just a matter of simply making 'good' decisions, but also depends on the existence of open and transparent decision making processes. Specifically, it is well established in the literature on state-building[1] that a primary source of legitimacy is the involvement in decision making by the local population (Chesterman 2004) or (unbiased) international institutions like the U.N. (Barnett 1997). Occupying or peace-keeping authorities that do not share

[1] But also on conflict resolution more generally, see Tyler and Lind (1988), or Tyler (2004).

or delegate power will sooner or later run into resistance from the population (Edelstein 2004). In other words, it matters not only "what" is being decided, but also "how" it is being decided.

From an economic perspective, this fact is puzzling. Economists typically assume that people are motivated by outcomes (the what), not by the processes that generate these outcomes (the how). As a consequence, economists have had little to say on issue of legitimacy, or the role of decision making procedures in state-building.

In this paper, I use a formal model to approach this puzzle. I show that even if people care only about outcomes (the what), procedures nevertheless matter if there is uncertainty about the intentions of the authority. In the context of a state building such uncertainty is especially important, since the population may fear, with reason, that foreign peace-keeping troops or occupiers aim to enrich themselves or exploit the territory for strategic reasons.[2] In these circumstances, open and transparent procedures will be a credible signal to the population that the authorities intentions are benign rather than exploitative. This raises the legitimacy of the occupier and the willingness of the population to cooperate, *independently* of the actual content of the decision making.

To generate this result, I formulate a simple signaling model between two players, an occupier or state-builder and a population. The starting point is uncertainty on the side of the occupied population about the 'type' of the occupier. The occupier can either be 'imperialistic' and care only about the personal benefit she derives from control of the territory, or she can be 'benevolent' and take the payoffs of the population into account to some extent. The model has two stages, a decision-making stage and a cooperation stage. In the decision-making stage, the two parties face a conflict of interest in deciding whether to implement a policy A, favored by the population, or B, favored by the occupier. One can interpret these policies as different ways of (re)building the country's institutions, for example the content of a new constitution. In this stage, the occupier decides on the degree of delegation or participation embodied in the decision making procedure, which determines the probability that the policy will be the one that is preferred by the population.

[2]This fact motivates the application to post-war state-building, although the model could be applied to conflict situations more widely (see section 5.2). Note that although I will use the term "occupation", referring to post-invasion state-building as it is taking place in Afghanistan and Iraq, the model can also capture both peace-keeping missions supported by international troops, as in Kosovo or East-Timor. Of course, these are rather different situations from a legal and military point of view, but they share the essential element of uncertainty about the intentions of the authorities which drives the theoretical results.

When the outcome of the procedure is determined and the policy is chosen, the players enter the cooperation stage. In this stage they simultaneously choose costly cooperation levels. The cooperation levels jointly determine the payoffs of both parties during the occupation period. A crucial assumption is that there is complementarity in the cooperation levels, so that parties are mutually dependent on each other's cooperation to make state building successful.

I show formally that there exists a unique separating equilibrium of the game in which the degree of delegation allowed by the occupier is a credible signal of the latter's type. In such an equilibrium, procedures affect cooperation by the population for two reasons. First, a participatory procedure is more likely to yield the kind of institution formation that the population prefers. This raises the population's motivation to cooperate. I call this the *outcome effect*, which does not depend on the signaling role of procedures. Second, participatory procedures increase cooperation *even* if they result in an inferior outcome (from the population's point of view). This *procedural effect* arises because a participatory procedure reveals the benevolent intentions of the occupier, which increases legitimacy and reduces the fear of 'exploitation' on the side of the population.

The model adds in several ways to the literature on military occupations and conflict more generally. Although signaling explanations have been used in international relations and international politics (Walsh 2007), this paper is to my knowledge the first to use a signaling argument to explain the relation between legitimacy, participatory procedures and the substance of decision making. Second, the model gives intuitive predictions about when procedures will matter most. It predicts that the size of procedural effects rises with the initial degree of uncertainty about the occupier's intentions. Moreover, a necessary condition for procedural effects is that institution formation is sufficiently contested. If this is not the case, handing over control is not sufficiently costly for the occupier to be a credible signal. A further condition is a sufficient degree of mutual dependence between the occupier and the population. The model also adds to the economic literature on conflict resolution, because it shows that even if people care only about outcomes, the decision making *process* matters.

In the last part of the paper I present empirical evidence for the relevance of this signaling explanation. First, it is supported by Edelstein (2004)'s extensive empirical analysis of 24 different military occupations in the post-Napoleonic era. One of the main findings of Edelstein's exercise is that the delegation of power is is crucial to convince the population of the occupied territory that they will not be exploited, and therefore an important condition for the eventual success of an occupation. Second, for concreteness and illustration, I will link my assumptions

and results throughout the paper to the events during the initial stages of the American-led occupation of Iraq (2003-2004). As I will explain in more detail in the last section, the occupation that followed was characterized by procedural and outcome effects that can be explained in terms of the model in Section 3.

The paper proceeds as follows. The next section develops a formal definition of the notion of a 'procedure' and 'delegation'. I will then explain the model. Section 6 provides a more elaborate discussion of the relation of the model to the evens in Iraq and also discusses some other potential applications.

2 An operational account of delegation

In this section, I will propose a formal, stylized definition of the degree of delegation or participation in decision making. This definition will allow a distinction between procedures and outcomes that forms the basis of the model in Section 3.

During occupations, occupiers have several options in choosing the degree of delegation to local authorities. On one end of the extreme, they can keep all power to themselves, which allows them to retain maximum control over the kind of institution formation that occurs during the transition period. On the other hand, they can set up advisory boards, governing councils, shadow parliaments and so on. Each of these institutions may exhibit different degrees of independence, either formally through the way the rules are written, or informally, through the strength of the connections that their members maintain with the occupier.

Thus, implementing delegation or participation is not a zero-one decision, but admits many grey shades. To distinguish between these shades in a formal way I will use the following stylized definition of delegation.

Definition 1 (Delegation) *A procedure implements 'delegation of degree q' if the ex-ante probability that the population obtains its preferred outcome through the procedure is q.*

This definition interprets procedures as stochastic processes, the outcome of which is uncertain ex-ante. However, the occupier *can* determine ex-ante how likely a given procedure is to provide outcomes that are in favor of the population in cases of conflicting preferences. One can interpret this probability as the degree to which control is delegated to the representatives of the population, or the bargaining power of these representatives.[3]

[3]This definition strongly relates to a formal literature in economics on the delegation of decision rights (Aghion

This definition is extremely stylized, and precludes the analysis of many interesting phenomena. First, the notion of procedures proposed here abstracts completely from all aspects that relate to the process of decision-making itself. For example, it does not take into account that populations may have preferences to express their opinion or the psychological benefits of self-determination (Lind *et al.* 1990). I also abstract from the possibility that "the force of the better argument" (Habermas 1990: 158-9) may resolve conflicts of interest. In the current setup, procedures do not resolve conflicts of interests, they are merely institutions for managing them.

Second, this definition - and indeed the model in this paper - does not take into account heterogeneity in the population. The 'population' is assumed to have homogeneous preferences and can therefore be represented as a single actor. This simplification excludes discussions of 'power sharing' between ethnic groups (Sisk 1996). However, it still allows fruitful analysis of the transfer of sovereignty from an 'occupier' to a 'population', which is the focus of this paper.

Finally, one could also interpret q as measuring the delegation of power to a third party, like the U.N. or an NGO. What matters in this case is the ex-ante probability that in conflict situations the third party will take decisions that favor the population.

3 Delegation as a signal

In this section, I present a model of costly signaling with two different players: an occupier or authority (she) and a 'population'. The game is one of imperfect information, since the occupier has private information about her own preferences, which are determined by a chance-move by "nature". As discussed above, the population is assumed to have homogenous preferences, and can therefore be represented as a single agent. The players are indicated by subscripts $i \in \{a, p\}$, where a stands for 'authority' or occupier, and p for 'population'. During the game, players implement one of two different 'policies', denoted by $n \in \{A, B\}$, over which preferences of the players diverge. These 'policies' can relate to cultural, geopolitical, military or religious elements of institution formation. For example, in the case of Iraq, preferences between Americans and Iraqis diverged on the issue of the post-war constitution. The Americans favored a liberal Western style constitution, whereas many Iraqis favored a more conservative Islamic-based one (Arato 2009).

and Tirole 1997, Aghion *et al.* 2002, 2004). Dessein (2006) uses the notion of "stochastic control rights" in the context of financial contracting. Sebald (2009) has used a similar definition of procedures in his innovative paper on procedural effects.

Timing. The timing of the model (summarized in Figure 1) is as follows:

1. Nature determines the type a of the occupier. With probability χ the occupier is 'imperialistic' and with probability $1-\chi$ she is 'benevolent' (the meaning of these terms is explained below).

2. *Procedural stage.*

   (a) The occupier decides on the degree of delegation $q \in [0,1]$ of the procedure.

   (b) Nature decides the outcome of the procedure. The population gets its preferred policy (A) with probability q, and the occupier gets her preferred policy (B) with probability $1-q$.

3. *Cooperation stage.* After having observed both q and the outcome n, the population and the occupier simultaneously decide their level of cooperation $c_i(n,q) \in \mathbb{R}_+$.

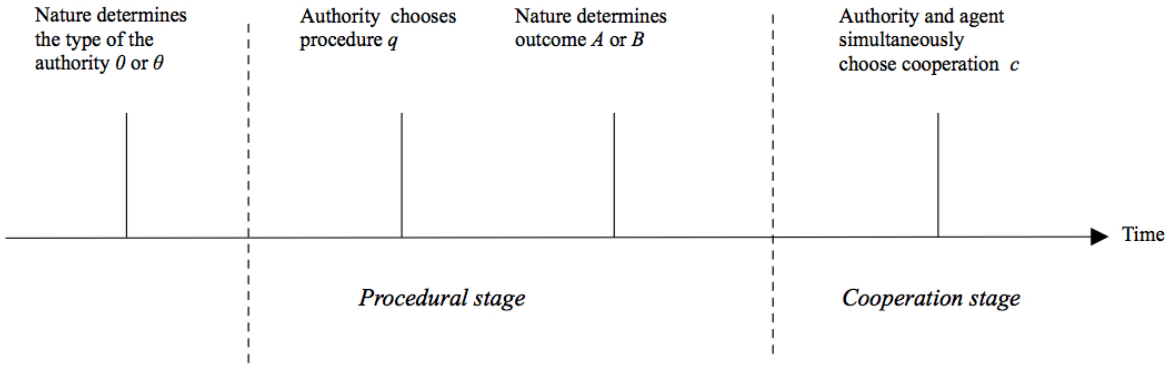


Figure 1: Timing of the game.

It is clear that the timing of the model is a simplification over the typically messy unfolding of occupations in the real world. Specifically, delegation will often occur gradually during a transition period (Chesterman 2004). Nevertheless, I believe that this structure captures the essential trade-offs of both players.

Payoffs from the policy. The benefit or output Y of the policy is determined by the cooperation levels of the two players as follows:

$$Y = c_a^{\gamma}\, c_p^{1-\gamma}, \tag{1}$$

where $\gamma \in (0,1)$ is a measure of the complementarity in cooperation levels.[4] If γ is close to 1, the occupier's productivity is not very much affected by the population's cooperation level. On the other hand, γ close to $1/2$ means that there is high complementarity between the cooperation choices, in the sense that a player's productivity depends highly on the cooperation of the other player. One can therefore interpret γ as *the degree of control* of the occupiers over the occupied territory. This may depend amongst other things on the amount of 'boots on the ground' and the size of the population.

'Cooperation' can be interpreted broadly as an activity that is both good for overall welfare and (as we will see below) individually costly. On the part of the occupier, it involves providing law, order and services to the population, as well as not exploiting the territory (for example by raiding the countries natural resources). On the part of the population it involves supplying labor, administrative cooperation and abstention from sabotage and resistance activities.

The payoff of each player depends on the type of policy. The population prefers policy A whereas the occupier prefers policy B. Note that it is not necessary that the occupier knows the preferences of the population with certainty. The qualitative results of the model go through if there is a *potential* conflict of interest, i.e., if the occupiers thinks that the population prefers policy A with some probability bigger than 0. The payoffs from the policy are:

$$\pi_i = \begin{cases} \lambda Y & \text{if policy } n \text{ is the preferred policy of player } i \\ (1-\lambda)Y & \text{if policy } n \text{ is not the preferred policy of player } i, \end{cases} \tag{2}$$

where $\lambda > \frac{1}{2}$ is the fraction of payoffs that accrue to player i. Thus, λ is a measure of the size of the *conflict of interest* between the two players. The larger is λ, the larger the difference in payoffs between the two policies, and the larger is the conflict of interest.

The population. The population has von Neumann-Morgenstern preferences given by

$$u_p = \pi_p - \frac{1}{2}c_p^2. \tag{3}$$

Here, the second term in the utility function captures the loss from the costly cooperation to the population. The quadratic shape of this function is standard, but any convex function

[4] Although more general formulations of a supermodular technology are conceivable, the Cobb-Douglas function has the advantage that it has a readily interpretable parameter that measures complementarity. Note also that I assume that complementarities in cooperation are part of the technology. However, such complementarities might arise from other sources as well. For example, if the players have reciprocal preferences, so that they cooperate more if the other has been 'nice', this will generate qualitatively similar results.

will generate similar results. As discussed above, the population is regarded as a single player, implying homogeneous preferences.

The occupier. The occupier has the following von Neumann-Morgenstern preferences

$$u_a = \pi_a + a\pi_p - \frac{1}{2}c_a^2. \tag{4}$$

Here a is determined by nature at the start of the game to be either 0 or θ with equal probability. Its realization is private information of the occupier. If $a = 0$ we call the occupier 'imperialist' or a low type, and u_0 denotes the utility of this type. One can think of such a type as being only interested in the occupied territory in so far as it can satisfy its geopolitical need or domestic demand for resources.

If $a = \theta \in (0, 1)$ we call the occupier 'benevolent' or a high type, and her utility is denoted by u_θ. If the occupier is benevolent, she has preferences over the payoffs π_p of the population from the chosen policy. Everything else equal, she prefers higher payoffs of the population from the policy. However, since $\theta < 1$ there is always a conflict of interest between the occupier and the population even if the occupier is benevolent, because she always values her own payoffs more. Note that since θ is private information of the occupier, one can also interpret the size of θ as a measure of the amount of uncertainty on the side of the population.

Throughout the paper, I will refer to the population's subjective ex-post belief $\mu \equiv Prob(a = \theta \mid q)$ that the occupier is benevolent as the *legitimacy* of the occupier. This fits well with the definition of legitimacy as "a subjective quality, relational between actor an institution and defined by the actors' perception of the institution"(Paris and Sisk 2009:15). It also fits the popular conception of legitimacy as an authority's right to govern.

In summary, the game consists of a simultaneous move game, preceded by a procedural stage that opens up the opportunity for the occupier to signal. Note that neither player has a preference over procedures q, they care only about the payoffs from the policy. This is in contrast to for example Frey *et al.* (2004) who aim to explain procedural effects by positing preferences over procedures. However, if such preferences were incorporated as a separable term it would not alter the qualitative results of this paper.

4 Signaling through delegation

In this section we look for a perfect Bayesian equilibrium of the game. As is customary in the literature on signaling games, we mandate that this equilibrium satisfies the Cho and Kreps (1987) 'Intuitive Criterion'. This criterion rules out equilibria that depend on off-equilibrium beliefs that are 'unreasonable' in a rather weak sense. All proofs are in the Appendix.

Let $c_i(n, q)$ denote the cooperation level of player i when procedure q resulted in outcome n, and denote equilibrium values by *. I start by analyzing the players' cooperation choices in the last round of the game. Because the payoffs are concave and the costs of cooperation are convex, the optimal cooperation exists and is bounded for each player. Denote by $R_i(n, q; c_{-i})$ the optimal cooperation of each player as a function of the cooperation of the other player, i.e., R_i is the reaction function of player i. The first important observation is that

$$R_\theta(n, q; c_p) \geq R_0(n, q; c_p), \tag{5}$$

i.e., for any given n and q, the benevolent occupier will always be more cooperative than the imperialistic occupier. The reason is that the benevolent occupier internalizes a part of the payoffs of the population, and therefore she is more motivated to exert cooperation.

This leads to the second observation, due to the fact that the cooperation levels are complements. Because the type of the occupier affects her cooperation level, beliefs about the occupier's type will affect cooperation of the population. If the latter believes the occupier is benevolent, it will be more motivated to exert high cooperation. Thus, interpreting the population's belief on the likelihood that the occupier is benevolent as the *legitimacy* of the occupier, a higher legitimacy translates into a higher cooperation level.

We now move to the procedural stage. Before characterizing the equilibrium level of q, we formulate a useful lemma, that will lay the basis for the existence of a separating equilibrium. Given that the occupier's cooperation in the second stage is a best response to the cooperation of the population, we can suppress some notation and write the expected utility of the occupier as follows:

$$U_a\left(q, c_p(n, q)\right) = q \cdot u_a\left(A, c_p(A, q)\right) + (1 - q) \cdot u_a\left(B, c_p(B, q)\right).$$

It is possible to show that this expected utility function satisfies the following single crossing condition

Lemma 1 *For any $q' > q$,*

$$U_0\left(q', c_p(n, q')\right) \geq U_0\left(q, c_p(n, q)\right) \Rightarrow U_\theta\left(q', c_p(n, q')\right) > U_\theta\left(q, c_p(n, q)\right). \tag{6}$$

Thus, whenever the imperialistic occupier is indifferent between a pair $(q', c_p(n, q'))$ and $(q, c_p(n, q))$, the benevolent occupier will strictly prefer the pair with the higher q.

Lemma 1 is known as a single-crossing condition, because it ensures that the indifference curves of the different types of occupiers cross at most once. The proof of this claim is algebraically somewhat involved but intuitively simple: raising q is less costly for the benevolent occupier than for the imperialistic occupier. Whereas the loss of her favorite policy is a pure loss to the latter, the benevolent occupier internalizes some of the gain to the population. It should be no surprise that a useful single crossing property in this game should be in *expected* utility. The signaling variable of the occupier is q, the probability with which her non-preferred outcome A occurs. Thus, although the signaling variable does not directly influence the occupier's outcome in any given policy, it changes her expected outcome through modifying the probability of each policy being chosen.

Outcome and procedural effects. Before we move on to the equilibrium results, I make an important distinction. A central point of the paper is to distinguish between two effects of the occupier's choices that determine the cooperation level of the population. The first is the effect of actual 'outcomes' (n) on cooperation levels, the second is the effect of the type of procedure (q). To avoid confusion, I provide formal definitions of both the outcome and the procedural effect.

Definition 2 *For a given level of participation $\bar{q}$, the outcome effect is $c_p(A, \bar{q}) - c_p(B, \bar{q})$.*

Definition 2 defines the 'outcome effect' as the increase in cooperation when the population's favorite policy (A) is chosen. Note that the term 'outcome' refers here to the outcome of the decision-making procedure (i.e., A or B), i.e. the kind of state-building that is pursued. It does not refer to the utility level of the population at the end of the game, or to the final success of the state-building exercise.

Definition 3 *For a given outcome $\bar{n} \in \{A, B\}$ and some values q and q' such that $q' > q$, the procedural effect is $c_p(\bar{n}, q') - c_p(\bar{n}, q)$.*

Definition 3 defines a procedural effect as the increase in cooperation due to a more participatory procedure.

With these definitions in hand, we can derive the existence and characteristics of a separating equilibrium in the game.

Proposition 1 (Separating equilibrium) *If and only if* $\theta \leq \left(\frac{\lambda}{1-\lambda}\right)^{\frac{3-\gamma}{1-\gamma}} - \left(\frac{1-\lambda}{\lambda}\right)$*, then there exists a unique separating equilibrium in pure strategies in which*

a) the imperialistic occupier chooses $q_0^* = 0$*,*

b) the benevolent occupier chooses

$$q_\theta^* = \begin{cases} q_\theta^*(\lambda, \theta, \gamma) < 1 & \text{if } \theta < \bar{\theta}(\lambda, \gamma) \\ 1 & \text{if } \theta \geq \bar{\theta}(\lambda, \gamma), \end{cases} \tag{7}$$

c) both the outcome and the procedural effect are positive, i.e.

$$c_p^*(B, q_0^*) < c_p^*(B, q_\theta^*) < c_p^*(A, q_\theta^*). \tag{8}$$

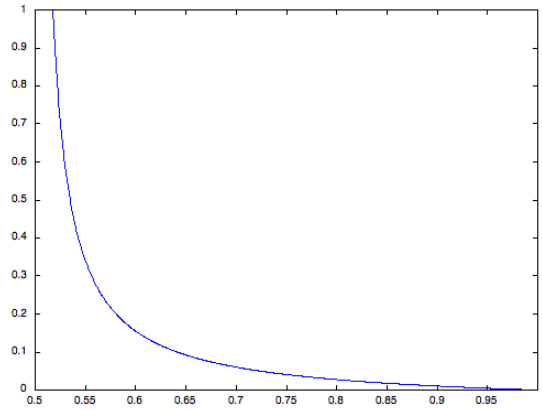
Proposition 1 says that the degree of delegation of the decision making procedure is a signal of the type of the occupier. Observing $q_\theta^* > 0$ means that the occupier is benevolent, whereas observing $q_0^* = 0$ means that the occupier is imperialistic.

To understand why this is an equilibrium, consider the occupier's choice between q_0^* and q_θ^*. From the point of view of the occupier, two things are relevant. On the one hand, choosing q_θ^* leads to higher equilibrium cooperation level by the population, as one can see from Proposition 1*c*). This increases utility to the occupier for a given outcome A or B. On the other hand, choosing a higher q increases the probability that the occupier will end up with the 'wrong' policy. The existence of the separating equilibrium stems from the fact that delegating control is more costly for the imperialistic occupier. The equilibrium exists only if the conflict of interest λ is high enough relative to θ. The intuition behind this condition is that a high θ increases the cooperation of the population when it observes q_θ^*, which makes it more profitable for the selfish type to mimic the signal. If λ is low relative to θ, the selfish type is willing to mimic even the strongest signal ($q = 1$) in which case the separating equilibrium collapses.

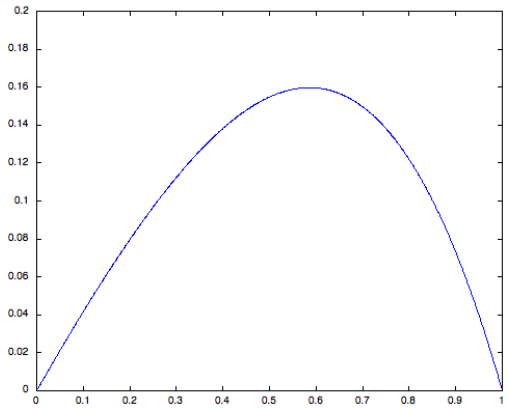
Proposition 1*b*) tells us that the procedure chosen by the benevolent occupier depends on θ, λ and γ. If θ is high relative to λ (but not so high that it violates the equilibrium condition), then the benevolent occupier prefers to implement policy A and sets $q_\theta^* = 1$. Note that the conflict of interest is still there: for given cooperation levels, the occupier would prefer to carry out policy

B. However, when θ is high, the increased cooperation of the population makes implementing A so attractive, that it outweighs the loss to the benevolent occupier of her favorite policy.

On the other hand, if θ is low relative to λ, the increased cooperation by the population does not compensate the loss of policy B. Then, the benevolent occupier sets q_θ^* at the lowest level such that the incentive compatibility constraint of the selfish type is satisfied, and the selfish type does not mimic the signal of the benevolent type. This incentive compatibility constraint is graphed in Figure 2, which depicts q_θ^* as a function of λ and γ. In the left panel we see that if λ increases, the optimal level of delegation drops. The reason is that if the conflict of interest increases, inducing participation becomes more costly. Thus, a lower q will be sufficient to deter the low type from copying the signal.



(a) q_θ^* as a function of λ, with $\theta, \gamma = 0.5$.



(b) q_θ^* as a function of γ, with $\theta = 0.5$, $\lambda = 0.6$.

Figure 2: The equilibrium q_θ^* as a function of the conflict of interest λ (panel a) and complementarity γ (panel b).

The right panel of Figure 2 shows how a change in γ affects q_θ^*. The humpshape occurs because complementarities are highest for intermediate values of γ. When complementarities are high, q_θ^* rises, because it now becomes more important to the occupier to insure cooperation from the population. This makes it more attractive for the low type to mimic the signal, and the high type will have to choose a stronger signal to distinguish herself.

Proposition 1c) is the main result of this paper, since it tells us that both the outcome and the procedural effect are positive. This means the model predicts that under the right conditions, both the content of decisions *and* the way they are taken matter for the cooperation of the

population with the occupiers. Note that this explanation of process effects relies entirely on conventional preferences. The reason is simply that delegation provides a credible signal of the occupier's preferences. In this way, they affect the population's expectations of the occupier's behavior (exploitative or cooperative) in the transition period.

Proposition 1c) also shows that a higher *legitimacy* of the occupier - the belief that she is a benevolent type - leads the population to be more cooperative. This is true because legitimacy is associated with better performance of the occupier, both in terms of the likely policy choice (outcome effect) and the actual cooperation level of the occupier for a given policy (procedural effect).

The uniqueness of the equilibrium results from applying the intuitive criterion to rule out pooling equilibria on low participation levels. In any pooling equilibrium, deviations to higher levels of q by the selfish type are dominated by the equilibrium strategy. As a result, the intuitive criterion rules out low off-equilibrium beliefs upon observing high values of q. This causes the benevolent type to deviate from any candidate pooling equilibrium.

Note that Proposition 1 also implies that the utility of the population is higher under participatory procedures. Because of the externalities that are present in the cooperative stage of the game, a Pareto improvement could be obtained if both populations would exert more than their equilibrium cooperation. A benevolent occupier takes (part of) this externality into account and this improves efficiency. Participatory procedures improve efficiency further by making the benevolence common knowledge, which in turn increases the cooperation of the population and the occupier. Thus, for a given outcome, utility of the population is higher under participatory procedures.

Proposition 1 shows that outcome and procedural effects can, and under some conditions will occur. But what determines the size of procedural and outcome effects? Proposition 2 shows how the parameters of the model affect the *size* of the procedural and outcome effects in the separating equilibrium:

Proposition 2 *In the separating equilibrium described in Proposition 1, the size of the procedural and outcome effect are a hump-shaped function of the degree of control γ. An increase benevolence θ and the conflict of interest λ affects the outcome and procedural effect as follows*

The intuitions behind the signs in the table in Proposition 2 are mostly straightforward. A higher degree of altruism θ raises the cooperation level of the benevolent authority. This increases the difference in the cooperation levels of the two types of authorities and therefore the size

	Procedural effect	Outcome effect
θ	$+$	$+$
λ	$-$	$+$

of the procedural effect. Here it is useful to think about θ as the degree of uncertainty of the population. The higher the uncertainty that is resolved in the separating equilibrium, the higher the procedural effect. A higher θ also causes the benevolent occupier to raise his cooperation more under policy A, since this is the policy that gives the population the highest utility. Thus, the population is more motivated to work on policy A and this causes an increase in the outcome effect. In sum, the more benevolent the authority potentially is, the more is at stake in resolving the uncertainty.

A rise in the conflict of interest λ diminishes the importance of procedures relative to outcomes. An increase in λ increases the motivation of the population to work on policy A and therefore increases the outcome effect. It also raises the cooperation of the occupier on policy B, which in turn raises the population's cooperation on policy B. Due to the decreasing returns to scale this diminishes the effect of revealing θ and therefore reduces the procedural effect. Thus, the model predicts that procedures are most important relative to outcomes for intermediately contested issues: for relatively uncontroversial issues, Proposition 1 shows that a separating equilibrium does not exist and hence there is no procedural effect, whereas for very controversial issues the outcome of the decision dominates the procedural effect.

5 Empirical relevance

The model predicts that the strength of the procedural effect depends on the amount of uncertainty about the type of the authority, the degree of mutual dependence of authority and the population and the size of the conflict of interest between the parties. All these conditions are typically present during a transitory post-war occupation, so one would expect procedural effects to occur. The problem is that signaling effects are notoriously difficult to identify, because they describe an invisible process: the interpretation and inference certain events by the population. In this section, I will discuss how the signaling explanation fits with the work of Edelstein (2004). To add concreteness to this analysis, I will relate the model to the events during the American led occupation in Iraq. Finally, I will show that procedural concerns come

up in other conflict-contexts with a similar structure.

Edelstein (2004)'s empirical analysis of 24 post-Napoleonic occupations supports the importance of the signaling effect of procedures for the success of military occupations. Like this paper, Edelstein highlights the trade off between control and delegation. In his words, successful occupations require striking "a balance between returning sovereignty to the occupied population and retaining sufficient control over the direction of the occupied territory." (Edelstein 2004: 69). He identifies several strategies to credibly signal these intentions. All of these strategies involve delegating decision making power in the short or intermediate term either to the indigenous population or to 'neutral' third parties like the U.N. These strategies are credible, because they carry a real cost to the occupier, in terms of a loss of control over the direction of the territory.[5]

Edelstein relates successful communication of intention to a metric of 'success' of an occupation, based (among other things) on the stability of the occupied territory and the cost of the occupation to the occupier. He shows that communicating benevolent intentions succeeded in only 12 out of 24 occupations. Of these 12, 7 were subsequently successful. In the 12 cases where the signal of eventual withdrawal was not credible, no occupations succeeded, two had mixed success, and 10 failed.

5.1 The occupation of Iraq

When the American-led coalition defeated Saddam Hussein on April 9, 2003, uncertainty amongst Iraqis about their motives was high. Although the Bush administration expressed its intention to liberate the Iraqi people, many people suspected that more imperialistic motives were underlying the invasion (Arato 2009: 22-23). A poll by the Iraq Centre for Research and Strategic Studies (ICRSS) shows that right after the invasion in May 2003, 44.6% of the Iraqi population saw the invaders as liberators or peacekeepers, while a roughly equal fraction of 47.4% saw them as occupiers. Thus, in terms of the model, θ was high.

Second, in some areas of institution formation there was a substantial conflict of interest λ between the occupiers and the Iraqis. For ideological, strategic and domestic political reasons, the Americans attached high importance to establishing Western style, secular democracy in Iraq that would be friendly to the U.S. However, the main ethnic fraction in Iraq consisted of

[5]By contrast, mere rhetoric on the part of the occupier is not credible. For example, during the first 25 years of their occupation of Egypt, the British made around a 120 official statements of their intentions to leave (Al-Sayyid 1968, p. xi), but ended up staying 72 years.

Shia Muslims, who preferred a more religiously oriented constitution and had strong ties with Iran. The U.S. thus feared that giving away decision power - for example over the design of the constitution - would result in an Islamic and Iran-friendly government (Arato 2009: 125).

Given the high uncertainty and conflicts of interest, the model therefore predicts that procedural concerns would be central in shaping public opinion and Iraqi cooperation. As it happened, the United States' administrative body in Iraqi, the Coalition Provisional Authority (CPA) headed by Paul Bremer, implemented virtually zero participation during the initial stages of the occupation. Arato writes that "instead of instituting a gradually more inclusive political process, the CPA did the reverse" (2009: 23). Cockburn (2006: 68) concurs that under the rule of the CPA, "The assurances about Iraqis ruling Iraq were soon forgotten. Everything was to be controlled by US advisers with a few British assistants. Iraqis were to play a secondary role." As a formal confirmation, on May 22, 2003, U.N. resolution 1483 gave the U.S. the status of occupiers of Iraq and the associated, open-ended legal mandate.[6]

The model predicts that as a result, trust in the intentions of the occupiers would drop and the cooperation levels would be low. Polling data confirm this prediction. A year after the invasion, in May 2004, an ICRSS poll shows that the number of people who saw the Americans as liberators or peacekeepers had dropped from 44.6% to 5%, while the number of people who saw the Americans as occupiers rose from 47.4% to 86.2%. A Worldpublicopinion poll in September 2006 found that most Iraqis were convinced that the U.S. were planning to use the country for its own strategic purposes: 77% of the Iraqis believed that the Americans were planning to establish permanent bases in Iraq, and only 18% believed that the U.S. would remove its forces once Iraq was stabilized. 76% thought that the Americans would not withdraw their forces within six months if the by then newly established Iraqi government asked them to do so. In addition, an increasing number of Iraqi people supported resistance against the occupation forces. A YouGov poll in July 2003 shows that 18% of the population was fairly to very hostile to the foreign forces, the rest being neutral or positive. By contrast, in January 2006, a poll by World Public Opinion found that 47% of Iraqi's approved of attacks on coalition forces.

There are several reasons to think that this drop in trust is at least partly attributable to procedural issues. First, some observers who were on the ground at the time cite a signaling

[6]One of the reasons for this strategy was that the Americans were confident of their ability to implement their plans by relying on their own military power. Cockburn (2006: 68-69) writes that "the war [against Saddam Hussein] had been so easy. The US thought it had no need for friends or allies." Thus, in terms of the model, the Americans thought they were facing a very high γ, and could operate virtually independently of Iraqi cooperation.

explanation. Galbraith writes

> "Every Iraqi leader, including the most pro-American, says it was Bremer's decision to keep power that changed the United States from being seen by many as liberator to being universally regarded as an occupier."
>
> Galbraith (2006: 122).

Steven Kull, director of the Program on International Policy Attitudes, argues that procedural effects are the main drivers of the hostility

> "It [the polls] suggests that what is key here is that the U.S. needs to address the feeling held by most Iraqis that they are not treated as a sovereign power. [...] As long as Iraqis do not have this sense, they are likely to continue to have a fundamentally hostile attitude toward all aspects of the U.S. presence in Iraq."
>
> Kull (2006)[7]

Second, the Americans themselves were aware of the symbolic effect of their authoritarian policies, and tried to *mimic* the existence of participatory decision making. Specifically, under pressure from the U.N. and prominent Iraqis, Bremer appointed the Interim Governing Council (IGC) in July 2003. The council was comprised of Iraqis from different religious and geographical backgrounds, and was, in the words of Bremer (2003), to "conduct the business of government". However, according to Diamond "The IGC was neither fish nor fowl: It was not really a 'governing' council, as Bremer made it clear that he would continue to exercise supreme power, including the power to veto any IGC decisions" (2005: 10). Arato describes the CPA as "[A] dictatorship that uses democratic public relations and a facade of participation, representation and consultation to hide its authoritarian practices." (Arato 2009: 18). In line with the model, this form of 'window-dressing' was not interpreted as a credible signal. In fact, Allawi (2007: 166) writes that Arab hostility was reinforced by "the deeply held conviction by the so-called Arab 'street' that the Governing Council was simply another instrument of control of the occupation."

Finally, there were individual instances where procedural concerns had a very explicit symbolic role. Perhaps the clearest case was the writing of a new Iraqi constitution. The Americans,

[7]Full testimony available at http//:www.worldpublicopinion.org/pipa/articles/home_aage/517.php

for fear of instituting an Islamic government, attempted to retain control over the constitution making process by appointing the members of the drafting committee. In response, prominent local clerics called for more democracy, and organized large demonstrations for a more democratic process. The American's backtracked, but only partially, and did control the eventual drafting process to a large degree (Arato 2009, Diamond 2005, Galbraith 2006).

Apart from procedural effects, the CPA also generated significant 'outcome effects' in its unilateral pursuit of the policy of reshaping Iraq in a Western image. It made sweeping 'Western-style' reforms (Fox 2005), some of which were contrary to Iraqi interests and had a detrimental impact on cooperation with the Americans.[8] As a result of both outcome and procedural effects, actual resistance against the coalition by violent groups rose steadily from 2003 through 2006 (Cockburn 2006, Allawi 2007). The result was an increasingly chaotic situation and rising death toll.[9]

Clearly, the model captures only some of the complex events in Iraq, and abstracts from important aspects such as the ethnic divisions in Iraqi society. Nevertheless, the models sheds light on the symbolic importance of procedures in the Iraq case, and can relate them to several fundamental characteristics of the Iraqi occupation.

5.2 Other applications

Procedural effects occur not only during military occupation, but in many areas of life where conflict is prevalent (e.g. Tyler and Lind 1992, Tyler 2004). The model presented in section 3 may be fruitfully applied to some of those areas. For example, in organizational contexts the output of the company will depend on the ability of management and the workers to cooperate constructively, so the complementarity between cooperation levels assumed in the model arises naturally. Moreover, management and employees may have different preferences over aspects of

[8]For example, the ambitious goal of the CPA to build a completely new government from scratch led it to disband the Iraqi army and large parts of the state apparatus. This resulted in firing of more than 400,000 army personnel and administrators, and in an economic crisis for about 10% of the population (Forman 2006: 204). This greatly facilitated recruitment for the insurgency (Ricks 2006: 191). The U.S. also imposed Western elements on the Iraqi constitution such as a bill of rights and a limited role of Islamic law (Galbraith 2006: 140, Arato 2009: chapter 3). This led some religious Iraqi leaders to reject cooperation with the Americans (Arato 2009, chapter 3).

[9]The organization Iraqi Body Count reports that "There were 6,332 reported civilian deaths in the 10.5 months following the initial invasion in year one, or 20 per day; 11,312 in year two, 55% up on year ones daily rate; 14,910 in year three (32% up on year two); and a staggering 26,540 in year four (78% up on year three, and averaging 74 per day)." Source: http://www.iraqbodycount.org/analysis/numbers/year-four

the workload and working conditions. There is indeed overwhelming evidence that procedural effects are prevalent in the workplace (Cohen-Charash and Spector 2004).

Another salient application is the relationship between politicians and citizens. For their welfare, citizens depend on politicians to work for the public good, and politicians depend on citizens for cooperation and compliance with the law. Frey (1997) and Frey *et al.* (2004) provide evidence that participatory decision making procedures like the use of referenda correlate with increased tax compliance.

6 Conclusion

Post-conflict state building is often characterized by substantial uncertainty and mistrust on the side of the population about the goals and intentions of the authority. This paper explains why in such a situation, decision processes are of great symbolic importance, and affect the legitimacy of the authorities independently from the content of decisions. Delegation leads to increased legitimacy because it signals the intentions of the occupier. One condition for this to happen is that there is enough at stake: if the conflict of interest is too small, delegating power is not a credible signal. There also needs to be mutual dependence between the parties in conflict: if the transitory authorities do not need cooperation from the population, signaling will not occur. Moreover, the larger the initial uncertainty, the larger is the effect of procedures.

In the model, people derive utility from final outcomes, and do not care for procures except for their informational value. This does not rule out that people may care about participation and delegation for more intrinsic reasons, such as a preference for expressing one's opinion, or the dignity that comes from being represented in the political arena. However, it shows that even in the absence of such concerns, procedures matter for conflict resolution.

References

Aghion, Philippe, Mathias Dewatripont, and Patrick Rey. 2002. "On Partial Contracting." *European Economic Review*, 46, 745-53.

Aghion, Philippe, Mathias Dewatripont, and Patrick Rey. 2004. "Transferable Control." *Journal of the European Economic Association*, 1:2, 115-38.

Aghion, Philippe and Jean Tirole. 1997. "Formal and Real Authority in Organizations." *Journal of Political Economy*, 105:1, 1-29.

Allawi, Ali A. 2007. *The occupation of Iraq: winning the war, losing the peace.* New Haven: Yale University Press.

Al-Sayyid, Afaf Lutfu. 1968. *Egypt and Cromer: A Study in Anglo-Egyptian Relations.* London: John Murray.

Arato, Andrew. 2009. *Constitution making under occupation. The politics of imposed revolution in Iraq.* New York: Columbia University Press. Participation," *Journal of the American Planning Association*, 35:4, 216-224.

Barnett, Michael N. 1997. "Review: Bringing in the New World Order: Liberalism, Legitimacy, and the United Nations." *World Politics*, 49:4, 526-51.

Bremer, Paul L. III. 2003. "Iraq's path to sovereignty." *The Washington Post*, Sept. 8, 2003.

Chesterman, S. 2004. *You, The People: The United Nations, Transitional Administration, and State-building.* Oxford: Oxford University Press.

Cockburn, Patrick. 2006. *The occupation: war and resistance in Iraq.* London: Verso.

Cohen-Charash, Yochi and Paul E. Spector. 2001. "The role of justice in organizations: a meta-analysis." *Organizational Behavior and Human Decision Processes*, 86, 278-321.

Dessein, Wouter. 2005. "Information and Control in Ventures and Alliances." *Journal of Finance*, 60:5, 2513-50.

Diamond, Larry. 2004. "Lessons from Iraq." *Journal of Democracy*, 16:1, 9-23.

Forman, Johanna Mendelson . 2006. "Striking out in Baghdad: How Postconflict Reconstruction Went Awry," in Fukuyama (ed.), *Nation-Building: Beyond Afghanistan and Iraq.* Baltimore: John Hopkins UP. 196-217.

Frey, Bruno S. 1997. "A Constitution for Knaves Crowds out Civic Virtues." *Economic Journal*, 107, 1043-53.

Frey, Bruno S., Matthias Benz and Alois Stutzer. 2004. "Introducing procedural utility: not only what, but also how matters." *Journal of Institutional and Theoretical Economics*, 160, 377-402.

Habermas, Jürgen. 1990. *Moral Consciousness and Communicative Action.* Cambridge: MIT Press.

Galbraith, Peter W. 2006. *The End of Iraq.* New York: Simon and Schuster.

Hurd, Ian. 1999. "Legitimacy and Authority in International Politics." *International Organization.* 53:2, 379 - 408

Lake, David A. 2009 *Hierarchy in International Relations.* Ithaca, NY: Cornell University Press.

Lind, Allan E. and Tom R. Tyler. 1988. *The Social Psychology of Procedural Justice.* New York and London: Plenum Press. Studies, 19, 17-38.

Paris, Roland and Timothy Sisk (eds). 2008. *Dilemmas of Statebuilding: Confronting the Contradictions of Postwar Peace Operations*, London: Routledge.

Sebald, Alexander. 2009. "Attribution and Reciprocity." *Games and Economic Behavior*, forthcoming.

Sisk, Timothy D. 1996. *Power Sharing and International Mediation in Ethnic Conflicts.* Washington DC: United States Peace Institute.

Tyler, Tom R. 2004. "Procedural Justice," in *The Blackwell Companion to Law and Society.* Austin Sarat (ed.) New York: Wiley Blackwell.

Tyler, Tom R. and Allan E. Lind. 1992. "A relational model of occupier in groups", in *Advances in Experimental Social Psychology (Vol. 25)* M. Zanna (ed.). New York: Academic.

Walsh, James I. 2007. "Do States Play Signaling Games?" *Cooperation and Conflict*, 42: 441 - 59.

Appendix: Proofs

Proof of Lemma 1. A sufficient condition for the result is that the indifference curves of the two types cross only once. This will guarantee that the change in c_p to compensate the occupier for a change in q is larger for the selfish occupier than for the benevolent occupier.

Let $\{(q, c_p(n,q;a))\}$ be the indifference curve of type a, i.e. $U_a(q, c_p(n,q;a)) \equiv \overline{U}_a$. Note that we have simplified notation by omitting $c_a(n,q)$ as an argument of U. This is without loss of generality: since c_a is set optimally in equilibrium, the envelope theorem tells us that its derivatives with respect to $c_p(n,q;a)$ are zero.

Thus, we want to show that for $n \in \{A, B\}$ the Spence-Mirrlees condition holds:

$$\frac{\partial c_p(n,q;0)}{\partial q} > \frac{\partial c_p(n,q;\theta)}{\partial q} \tag{A.1}$$

Taking the derivative of $U_a(q, c_p(n,q;a))$ w.r.t. q yields

$$0 = \frac{\partial U_a(q, c_p(n,q;a))}{\partial q} + \frac{\partial U_a(q, c_p(n,q;a))}{\partial c_p(n,q;a)} \frac{\partial c_p(n,q;a))}{\partial q}$$

$$0 = u_a\left(A, c_p(A,q;a)\right) - u_a\left(B, c_p(B,q;a)\right) + q\frac{\partial u_a\left(A, c_p(A,q;a)\right)}{\partial c_p(A,q;a)} \frac{\partial c_p(A,q;a)}{\partial q} + (1-q)\frac{\partial u_a\left(B, c_s(B,q;a)\right)}{\partial c_p(B,q;a)} \frac{\partial c_p(B,q;a)}{\partial q}$$

For simplicity, let us first investigate the case where $\partial c_p(B,q;a)/\partial q = 0$. Then we can write:

$$\frac{\partial c_p(A,q;a)}{\partial q} = \frac{u_p\left(B, c_p(B,q;a)\right) - u_a\left(A, c_p(A,q;a)\right)}{q\frac{\partial u_a(A,c_p(A,q;a))}{\partial c_p(A,q;a)}}. \tag{A.2}$$

It remains to derive the components of (A.2) and show that (A.1) holds. First consider the denominator. Using the envelope theorem as outlined above, it simplifies to:

$$\frac{\partial u_a\left(A, c_p(A,q;a)\right)}{\partial c_p(A,q;a)} = \frac{\partial(\pi_a + p\pi_p)}{\partial c_p(A,q;a)}.$$

To ease notation we denote $c_i(n,q;a)$ by c_{in}. We find

$$\frac{\partial \pi_0\left(A, c_{pA}\right)}{\partial c_{pA}} = (1-\lambda)(1-\gamma)\left(c_{pA}\right)^{-\gamma}\left(c_{aA}\right)^{\gamma}, \text{ and} \tag{A.3}$$

$$\frac{\partial(\pi_\theta\left(A, c_{pA}\right) + \theta\pi_p\left(A, c_{pA}\right))}{\partial c_{pA}} = (1-\lambda+\theta\lambda)(1-\gamma)\left(c_{pA}\right)^{-\gamma}\left(c_{aA}\right)^{\gamma} \tag{A.4}$$

Now we specify the numerator of (A.2). Denote $\Delta_a = u_a\left(B, c_p(B,q;a)\right) - u_a\left(A, c_p(A,q;a)\right)$. We have

$$\Delta_0 = \left[\lambda\left(c_{pB}\right)^{1-\gamma}\left(c_{0B}\right)^{\gamma} - \frac{\left(c_{0B}\right)^2}{2}\right] - \left[(1-\lambda)\left(c_{pA}\right)^{1-\gamma}\left(c_{0A}\right)^{\gamma} - \frac{\left(c_{0B}\right)^2}{2}\right] \tag{A.5}$$

$$\Delta_\theta = \left[(\lambda+(1-\lambda)\theta)\left(c_{pB}\right)^{1-\gamma}\left(c_{\theta B}\right)^{\gamma} - \frac{\left(c_{\theta B}\right)^2}{2}\right] - \left[(1-\lambda+\theta\lambda)\left(c_{pA}\right)^{1-\gamma}\left(c_{\theta A}\right)^{\gamma} - \frac{\left(c_{\theta B}\right)^2}{2}\right] \tag{A.6}$$

The best responses of the occupier $R_a(n,q;c_{pn})$ in the simultaneous move game are readily calculated by taking first order conditions from the occupier's' payoffs:

$$R_0(A,q;c_{pA}) = \left[(1-\lambda)\gamma\left(c_{pA}\right)^{1-\gamma}\right]^{\frac{1}{2-\gamma}} \tag{A.7}$$

$$R_0(B,q;c_{pB}) = \left[\lambda\gamma\left(c_{pB}\right)^{1-\gamma}\right]^{\frac{1}{2-\gamma}} \tag{A.8}$$

$$R_\theta(A,q;c_{pA}) = \left[(1-\lambda+\theta\lambda)\gamma\left(c_{pA}\right)^{1-\gamma}\right]^{\frac{1}{2-\gamma}} \tag{A.9}$$

$$R_\theta(B,q;c_{pB}) = \left[(\lambda+(1-\lambda)\theta)\gamma\left(c_{pB}\right)^{1-\gamma}\right]^{\frac{1}{2-\gamma}} \tag{A.10}$$

We insert these best responses of the occupier into the expressions for the denominator (A.3), (A.4) and the numerator (A.5), (A.6). After some tedious algebra we can define

$Z = \left[\gamma^{\frac{\gamma}{2-\gamma}} \left(c_{pB}\right)^{\frac{(1-\gamma)^2}{2-\gamma}} - \gamma^{\frac{2}{2-\gamma}} \left(c_{pB}\right)^{\frac{2(1-\gamma)}{2-\gamma}}\right]$ and $Q = -c_{pA} + \gamma^{\frac{2}{2-\gamma}} \left(c_{pA}\right)^{\frac{2(1-\gamma)}{2-\gamma}}$ and write:

$$\frac{\partial c_p(A,q;0)}{\partial q} > \frac{\partial c_p(A,q;\theta)}{\partial q}$$

$$\left(\frac{\lambda}{1-\lambda}\right)^{\frac{2}{2-\gamma}} Z + Q > \left(\frac{\lambda + (1-\lambda)\theta}{1-\lambda+\theta\lambda}\right)^{\frac{2}{2-\gamma}} Z + Q$$

$$\lambda > \frac{1}{2}$$

which is always satisfied.

An analogue argument exists if $\partial c_p(B,q;a)/\partial q > 0$ and $\partial c_p(A,q;a)/\partial q = 0$, but this is omitted here for reasons of space. It follows that the single crossing property holds if $\partial c_p(n,q;a)/\partial q > 0$ for all n. Thus, we showed that for a small increase in q, the high type needs to be compensated with an increase in $c_p(n,q)$ which is smaller than for the low type. This means the indifference curves cross only once. ■

Proof of Proposition 1. The outline of the proof is as follows. I first derive the optimal levels of cooperation of the occupier and the population. Then I show the conditions under which exists a $\tilde{q} > 0$ with associated beliefs $\mu\left(p = \theta \mid \tilde{q}\right) = 1$, such that the selfish occupier is indifferent between setting $\tilde{q}$ and $q = 0$ (the equilibrium level in the separating equilibrium). I then show that the benevolent occupier always prefers to set $\tilde{q}$ to any other q. Finally, I prove uniqueness by ruling out pooling equilibria. To ease notation in the remainder I denote cooperation levels $c_i(n,q)$ by c_{in}.

Equilibrium cooperation levels. Optimal cooperation levels are derived by taking derivatives of π_i for the population and both types of occupiers, which yields the reaction functions. Equilibrium levels are then computed by solving the system of reaction functions for cooperation levels. Note that cooperation levels will depend on the information of the population. It is straightforward to compute the equilibrium cooperation levels under complete information (as is the case in a separating equilibrium). The optimal cooperation level of the population is

$$c^*_{pA} = \left[\lambda\left(1-\gamma\right)\right]^{\frac{2-\gamma}{2}} \left[\gamma(1-\lambda+1_\theta\theta\lambda)\right]^{\frac{\gamma}{2}} \tag{A.11}$$

$$c^*_{pB} = \left[(1-\lambda)(1-\gamma)\right]^{\frac{2-\gamma}{2}} \left[\gamma(\lambda+1_\theta(1-\lambda)\theta)\right]^{\frac{\gamma}{2}} \tag{A.12}$$

Where $1_\theta = 1$ if the population beliefs the occupier is a high type and 0 if she beliefs he is a low type. We will denote by $\bar{c}_{pn}$ the optimal cooperation level of the population under maximally optimistic beliefs and by $\underline{c}_{pn}$ the optimal cooperation level of the population under maximally

pessimistic beliefs. The equilibrium cooperation level of the occupier is:

$$c_{aA}^{*} = \left[\lambda\left(1-\gamma\right)\right]^{\frac{1-\gamma}{2}}\left[\gamma(1-\lambda+1_{\theta}\theta\lambda)\right]^{\frac{1+\gamma}{2}} \tag{A.13}$$

$$c_{aB}^{*} = \left[(1-\lambda)(1-\gamma)\right]^{\frac{1-\gamma}{2}}\left[\gamma(\lambda+1_{\theta}(1-\lambda)\theta)\right]^{\frac{1+\gamma}{2}} \tag{A.14}$$

Note that when the we calculate the deviations to off-equilibrium actions, the beliefs of the population may not be correct. Specifically, when the population (mistakenly) thinks the occupier is a high type, the optimal cooperations of the selfish occupier are given by

$$\hat{c}_{0A}^{*} = \left[\lambda\left(1-\gamma\right)\right]^{\frac{1-\gamma}{2}}\left[(1-\lambda)\gamma\right]^{\frac{1+\gamma}{2}}\left[1+\theta\left(\frac{\lambda}{1-\lambda}\right)\right]^{\frac{\gamma(1-\gamma)}{2(2-\gamma)}} \tag{A.15}$$

$$\hat{c}_{0B}^{*} = \left[(1-\lambda)(1-\gamma)\right]^{\frac{1-\gamma}{2}}\left[\gamma\lambda\right]^{\frac{1+\gamma}{2}}\left[1+\theta\left(\frac{1-\lambda}{\lambda}\right)\right]^{\frac{\gamma(1-\gamma)}{2(2-\gamma)}} \tag{A.16}$$

Indifference of the selfish type. Assume there is a $\bar{q}>0$ with associated beliefs $\mu\left(p=\theta \mid \bar{q}\right)=1$. Similarly, assume that $\mu\left(p=0 \mid q=0\right)=1$. Then, it is easy to see from the optimal cooperation levels derived above that $\underline{c}_{pn}<\bar{c}_{pn}$.

A necessary condition for the existence of a separating equilibrium is that $u_0\left(B,\underline{c}_{pB}\right)\geq u_0\left(A,\bar{c}_{pA}\right)$, where we suppress the best response cooperation of the occupier as a part of the utility function to save notation. If this is not satisfied, the selfish occupier would always like to mimic the high type. This condition implies

$$\begin{aligned}
u_0\left(A,\bar{c}_{pA}\right) &\leq u_0\left(B,\underline{c}_{pB}\right)\\
(1-\lambda)Y\left(\bar{c}_{pA},\hat{c}_{0A}^{*}\right)-\frac{1}{2}\left(\hat{c}_{0A}^{*}\right)^{2} &\leq \lambda Y\left(\underline{c}_{pB},c_{0B}^{*}\right)-\frac{1}{2}\left(c_{0B}^{*}\right)^{2}\\
(1-\lambda)\left[1-\frac{\gamma}{2}\right]\left[\lambda(1-\gamma)\right]^{1-\gamma}\left[\gamma(1-\lambda)\right]^{\gamma}\left[1+\theta\left(\frac{\lambda}{1-\lambda}\right)\right]^{\frac{\gamma(1-\gamma)}{2-\gamma}} &\leq \lambda\left[1-\frac{\gamma}{2}\right]\left[(1-\lambda)(1-\gamma)\right]^{1-\gamma}\left[\gamma\lambda\right]^{\gamma}\\
\left[1+\theta\left(\frac{\lambda}{1-\lambda}\right)\right]^{\frac{\gamma(1-\gamma)}{2-\gamma}} &\leq \left(\frac{\lambda}{1-\lambda}\right)^{2\gamma}\\
\theta &\leq \left(\frac{\lambda}{1-\lambda}\right)^{\frac{3-\gamma}{1-\gamma}}-\left(\frac{1-\lambda}{\lambda}\right)\equiv\hat{\theta}(\lambda,\gamma)
\end{aligned} \tag{A.17}$$

We know that $u_0(B,\underline{c}_{pB})<u_0(B,\bar{c}_{pB})$. Then, if (A.17) holds, by the continuity of U_0 in q, there exists a $\tilde{q}$ such that the selfish type is indifferent between $\{\tilde{q},\bar{c}_{pn}\}$ and $\{0,\underline{c}_{pn}\}$. Furthermore, because $u_0(A,\bar{c}_{pA})<u_0(B,\bar{c}_{pB})$ and U_0 is linear in q, it follows that the selfish type always prefers $q=0$ to any $\bar{q}\in[\tilde{q},1]$. (I follow the tiebreaking rule that an indifferent occupier chooses the lowest q). If (off-equilibrium) beliefs $\mu\left(p=\theta \mid q<\tilde{q}\right)$ are low enough, she will also not deviate to $p\in(0,\tilde{p})$.

Strategy of the benevolent type. By Lemma 1 we know that the high type prefers to set $\tilde{q}$ to $q = 0$, so $\tilde{q}$ is a candidate for a separating equilibrium. What about deviations to other levels of q? Consider first deviations to $\hat{q} \in (0, \tilde{q})$. Suppose off equilibrium beliefs are such that $\mu(p = \theta \mid \hat{q}) = 0$. These beliefs do not violate the Intuitive Criterion, because we know from (A.17) that the selfish type would always mimic the benevolent type if she were to set $q \in (0, \tilde{q})$ and beliefs are such that $\mu(p = \theta \mid \hat{q}) = 1$. Then, because U_θ is linear in q, $U_\theta(\hat{q}) < U_\theta(q = 0) < U_\theta(\tilde{q})$. Thus, such deviations are not profitable.

Now consider deviations to $\hat{q} \in (\tilde{q}, 1]$. Because a low type will never set q in this interval, the intuitive criterion tells us that the only reasonable off-equilibrium beliefs are $\mu(p = \theta \mid \hat{q} \geq \tilde{q}) = 1$. Thus, deviations are profitable if:

$$u_\theta(A, \bar{c}_{pA}) \geq u_\theta(B, \bar{c}_{pB})$$

$$(1 - \lambda + \theta\lambda) Y(\bar{c}_{pA}, c^*_{\theta A}) - \frac{1}{2}(c^*_{\theta A})^2 \geq (\lambda + (1-\lambda)\theta) Y(\bar{c}_{pB}, c^*_{\theta B}) - \frac{1}{2}(c^*_{\theta B})^2$$

$$\left[(1 - \lambda + \lambda\theta)\left(1 - \frac{\gamma}{2}\right)\right][\lambda(1-\gamma)]^{1-\gamma}[\gamma(1 - \lambda + \theta\lambda)]^{\gamma} \geq \left[(\lambda + (1-\lambda)\theta)\left(1 - \frac{\gamma}{2}\right)\right][(1-\lambda)(1-\gamma)]^{1-\gamma}[\gamma(\lambda + 1 - \lambda)\theta)]^{\gamma}$$

$$\left(\frac{\lambda}{1-\lambda}\right)^{1-\gamma} \geq \left(\frac{\lambda + (1-\lambda)\theta}{1 - \lambda + \lambda\theta}\right)^{1+\gamma}$$

$$\theta \geq \left(\frac{\lambda - (1-\lambda)\left(\frac{\lambda}{1-\lambda}\right)^{\frac{1-\gamma}{1+\gamma}}}{\lambda\left(\frac{\lambda}{1-\lambda}\right)^{\frac{1-\gamma}{1+\gamma}} - (1-\lambda)}\right) \equiv \bar{\theta}(\lambda, \gamma) \quad \text{(A.18)}$$

Suppose first that (A.18) holds. In that case, the high type is better off under policy A then under policy B. Thus, she will set $q^*_\theta = 1$.

Now suppose that (A.18) does not hold. In this case the high type prefers to implement policy B. Because U_θ is linear in q, the high type will never deviate to $\bar{q} \in (\tilde{q}, 1]$. It follows that there is a separating equilibrium in which the low type sets $q = 0$ and the high type sets $q^*_\theta = \tilde{q}$, where $\tilde{q}$ is such that it satisfies the incentive constraint of the low type

$$u_0\left(B, \underline{c}_{pB}\right) = q^*_\theta * u_0(A, \bar{c}_{pA}) + (1 - q^*_\theta) * u_0(B, \bar{c}_{pB}), \text{ or}$$

$$q^*_\theta = \frac{u_0\left(B, \underline{c}_{pB}\right) - u_0(B, \bar{c}_{pB})}{u_0(A, \bar{c}_{pA}) - u_0(B, \bar{c}_{pB})}$$

Substituting in the appropriate expressions yields

$$q^*_\theta = \tilde{q} = \frac{\lambda - \lambda\left[1 + \theta\left(\frac{1-\lambda}{\lambda}\right)\right]^{\frac{\gamma(1-\gamma)}{2-\gamma}}}{(1-\lambda)\left[1 + \theta\left(\frac{\lambda}{1-\lambda}\right)\right]^{\frac{\gamma(1-\gamma)}{2-\gamma}} - \lambda\left[1 + \theta\left(\frac{1-\lambda}{\lambda}\right)\right]^{\frac{\gamma(1-\gamma)}{2-\gamma}}} > 0. \quad \text{(A.19)}$$

It remains to show that both these high type strategies can occur in equilibrium, which is the case if $\hat{\theta}(\lambda,\gamma)$ derived in (A.17) and $\bar{\theta}(\lambda,\gamma)$ derived in (A.18)cross at least once. It is easy to derive that both $\hat{\theta}(\lambda,\gamma)$ and $\bar{\theta}(\lambda,\gamma)$ are monotonic and that

$$\lim_{\lambda\downarrow\frac{1}{2}}\left(\left(\frac{\lambda}{1-\lambda}\right)^{\frac{3-\gamma}{1-\gamma}}-\left(\frac{1-\lambda}{\lambda}\right)\right)=0 \qquad \text{and} \quad \lim_{\lambda\uparrow 1}\left(\left(\frac{\lambda}{1-\lambda}\right)^{\frac{3-\gamma}{1-\gamma}}-\left(\frac{1-\lambda}{\lambda}\right)\right)=\infty$$

$$\lim_{\lambda\downarrow\frac{1}{2}}\left(\frac{\lambda-(1-\lambda)\left(\frac{\lambda}{1-\lambda}\right)^{\frac{1-\gamma}{1+\gamma}}}{\lambda\left(\frac{\lambda}{1-\lambda}\right)^{\frac{1-\gamma}{1+\gamma}}-(1-\lambda)}\right)=\gamma \qquad \text{and} \quad \lim_{\lambda\uparrow 1}\left(\frac{\lambda-(1-\lambda)\left(\frac{\lambda}{1-\lambda}\right)^{\frac{1-\gamma}{1+\gamma}}}{\lambda\left(\frac{\lambda}{1-\lambda}\right)^{\frac{1-\gamma}{1+\gamma}}-(1-\lambda)}\right)=0$$

So, $\bar{\theta}(\lambda,\gamma)$ crosses $\hat{\theta}(\lambda,\gamma)$ only once and from above.

Uniqueness. Consider now a candidate pooling equilibrium on some level $0 < q_{pool} < \tilde{q}$ and some level of cooperation of the population $c_{pn}(q_{pool})$ supported by beliefs $\mu(a=\theta \mid q_{pool}) = E[p]$. This would indeed be an equilibrium if it were supported by sufficiently low off-equilibrium beliefs. However, we can rule out such beliefs for deviations to high levels of q by applying the intuitive criterion.

We know that in this pooling equilibrium the low type has a higher utility than in the separating equilibrium above (or she would deviate to $q=0$). This implies that if (A.17) holds, there exists a $\hat{q} < \tilde{q}$ such that the low type is indifferent between $(q_{pool}, c_{pn}(q_{pool}))$ and $(\hat{q}, \bar{c}_{pn})$, where, as before, $\bar{c}_{pn}$ is the optimal cooperation level of the population under maximally optimistic beliefs. Because the expected utility of the low type is continuously decreasing in q, deviating to any $q > \hat{q}$ is dominated for the low type, and as a consequence the intuitive criterion prescribes that $\mu(a=\theta \mid q \geq \hat{q}) = 1$. By Lemma 1 we know that if the low type is indifferent between $(q_{pool}, c_{pn}(E))$ and $(\hat{q}, \bar{c}_{pn})$, the high type prefers to set $(\hat{q}, \bar{c}_{pn})$, and therefore prefers to deviate. This implies that the candidate pooling equilibrium is not an equilibrium. ■

Proof of Proposition 2. Following Definitions 3 and using the notation developed above, the outcome effect OE is given by

$$\begin{aligned} OE &= \bar{c}_{pA} - \bar{c}_{pB} \\ &= (1-\gamma)^{\frac{2-\gamma}{2}}\gamma^{\frac{\gamma}{2}}\left[\lambda^{\frac{2-\gamma}{2}}(1-\lambda+\theta\lambda)^{\frac{\gamma}{2}} - (1-\lambda)^{\frac{2-\gamma}{2}}(\lambda+(1-\lambda)\theta)^{\frac{\gamma}{2}}\right] \end{aligned} \tag{A.20}$$

where we substituted the equilibrium cooperation levels from (A.11) and (A.12). Similarly,

following Definition 3, the procedural effect PE is given by

$$\begin{aligned} PE &= \bar{c}_{pB} - \underline{c}_{pB} \\ &= \left[(1-\lambda)(1-\gamma)\right]^{\frac{2-\gamma}{2}} \left[\lambda\gamma\right]^{\frac{\gamma}{2}} \left[\left(1+\theta\left(\frac{1-\lambda}{\lambda}\right)\right)^{\frac{\gamma}{2}} - 1\right] \end{aligned} \tag{A.21}$$

Taking derivatives (omitted here for reasons of space) yields the first order conditions with the signs described in Proposition 2. ■