A PRINCIPAL-AGENT PERSPECTIVE ON COUNTERINSURGENCY SITUATIONS

by

Federico S. Perez Duenas

June 2011

Thesis Advisor: Roberto Szechtmn
Second Reader: Michael Atkinson

Approved for public release; distribution is unlimited
A Principal-Agent Perspective on Counterinsurgency Situations

Federico S. Perez Duenas

Naval Postgraduate School
Monterey, CA 93943

Approved for public release; distribution is unlimited

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. I.R.B. Protocol Number: N.A.

The successful conclusion to the war in Afghanistan necessitates an array of solutions that includes non-kinetic actions. In particular, there is growing interest in having the local population take control of the regional security, providing protection from external threats as well as internal sources of violence. In this thesis, the principal-agent paradigm is used to model the relationship between coalition forces (principal) and local tribes (agents). These relationships are affected by the lack of alignment of interests as well as the principal's lack of knowledge about the actions taken by the agents. Two scenarios are considered: The simplest where the agents do not have private information and their actions can be verified. This scenario is used to determine a baseline for the transfers offered by the principal. The second — and more realistic — scenario captures the principal’s lack of knowledge about the actions taken by the agents. The last model provides qualitative insights about the cost to the principal due to unknown information.
A PRINCIPAL-AGENT PERSPECTIVE ON COUNTERINSURGENCY SITUATIONS

Federico S. Perez Duenas
Lieutenant Commander, Spanish Navy
B.S. degree in Mathematics, UNED-Spain, 2008

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2011

Author: Federico S. Perez Duenas

Approved by: Roberto Szechtman
Thesis Advisor

Michael Atkinson
Second Reader

Robert F. Dell
Chair, Department of Operations Research
THIS PAGE INTENTIONALLY LEFT BLANK
ABSTRACT

The successful conclusion to the war in Afghanistan necessitates an array of solutions that includes non-kinetic actions. In particular, there is growing interest in having the local population take control of the regional security, providing protection from external threats as well as internal sources of violence. In this thesis, the principal-agent paradigm is used to model the relationship between coalition forces (principal) and local tribes (agents). These relationships are affected by the lack of alignment of interests of the principal and the agents. To achieve the desired alignment, the principal must provide incentives to the agents. Two scenarios are considered: The simplest where the agents do not have private information and their actions can be verified. This scenario is used to determine a baseline for the transfers offered by the principal. The second — and more realistic — scenario captures the principal’s lack of knowledge about the actions taken by the agents. The last model provides qualitative insights about the cost to the principal due to unknown information.
# Table of Contents

## 1 Introduction
1.1 Introduction ........................................... 1
1.2 Motivation and Research Focus. .......................... 3
1.3 Contributions of This Work ................................ 5
1.4 Structure of the Thesis .................................. 5

## 2 Background
2.1 Socio-Political Context .................................. 7
2.2 Agency Background ....................................... 8

## 3 Modeling Framework
3.1 Basic Situation .......................................... 15
3.2 Violence Generation ...................................... 15
3.3 The Principal ........................................... 17
3.4 The Agents ............................................... 18

## 4 Reservation Utility

## 5 Analysis
5.1 The Case of Full Information ............................ 25
5.2 Problem With Partial Information ...................... 28

## 6 Conclusions and Extensions
6.1 Motivation and Setting .................................. 35
6.2 Analysis .................................................. 35
6.3 Extensions ................................................ 36

List of References ............................................. 39
Appendix 41

Initial Distribution List 43
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Basic principal-agent model.</td>
<td>4</td>
</tr>
<tr>
<td>3.1</td>
<td>Dependence among key variables.</td>
<td>16</td>
</tr>
<tr>
<td>3.2</td>
<td>Distribution of the violence from agent A.</td>
<td>18</td>
</tr>
<tr>
<td>5.1</td>
<td>Time line of the contractual game.</td>
<td>25</td>
</tr>
<tr>
<td>5.2</td>
<td>Evolution of the agents’ utilities with the transfers.</td>
<td>29</td>
</tr>
<tr>
<td>5.3</td>
<td>Time line of the contractual game.</td>
<td>29</td>
</tr>
<tr>
<td>5.4</td>
<td>Second best transfers with moral hazard.</td>
<td>31</td>
</tr>
<tr>
<td>A.1</td>
<td>Afghanistan Ethno Linguistic Map from the Institute for the Study of War.</td>
<td>41</td>
</tr>
</tbody>
</table>
ACRONYMS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANA</td>
<td>Afghan National Army</td>
</tr>
<tr>
<td>ANP</td>
<td>Afghan National Police</td>
</tr>
<tr>
<td>HIPC</td>
<td>Heavily Indebted Poor Countries</td>
</tr>
<tr>
<td>KKT</td>
<td>Karush-Kuhn-Tucker</td>
</tr>
<tr>
<td>MOF</td>
<td>Ministry of Finance</td>
</tr>
<tr>
<td>R2P</td>
<td>Responsibility to Protect</td>
</tr>
</tbody>
</table>
THIS PAGE INTENTIONALLY LEFT BLANK
EXECUTIVE SUMMARY

Conflicts between different groups where at least one group tries to influence the actions of the others have been widely studied in recent years. Here, the principal-agent model is used to analyze a situation comprising several stakeholders. In this model there is a reference player, the principal, who takes certain actions aimed at reducing the level of violence. The other players, the agents, decide whether to cooperate with the principal depending on the incentives offered by the principal — if their current situation improves — or if they simply do not have any other options.

This thesis is motivated by the current situation in Afghanistan, which began in 2001 when the Taliban were routed from Kabul, retreating toward the mountainous border region between Afghanistan and Pakistan. Since then, the United States government and its allies have attempted to stabilize Afghanistan. Here, the situation faced by coalition forces in many small towns and villages of Afghanistan is analyzed. These coalition forces are routinely exposed to attacks from insurgents, often with active participation or help from the local population. These communities tend to be motivated by self-interest and self-sufficiency, preferring to secure their own villages rather than have outsiders do it for them.

It is now widely accepted that a successful conclusion to the war in Afghanistan necessitates an array of solutions that includes non-kinetic actions, especially those in which the local population plays a key role. In particular, there is growing interest in having the local population take control of the regional security, providing protection from external threats as well as internal sources of violence.

In this thesis, the principal-agent paradigm is used to model the relationship between coalition forces (principal) and local tribes (agents). These relationships are affected by the following issues: The principal and the agents may not have aligned preferences and, without incentives, the agents will follow the interest of their shura. To achieve the desired alignment, the principal must provide incentives to the agents to make cooperation sufficiently rewarding. The agents may take actions that cannot be observed by the principal. For example, the principal may not know the extent of the agents’ cooperation with insurgent groups, or the effort put forth by the agents in curtailing violence (moral hazard).

The main contribution of this thesis is the introduction of a model describing what occurs when
government forces attempt to keep lower violence organically from the local tribes, in a counterinsurgency situation. Two sources of violence are considered: endogenous and exogenous. The endogenous violence is due to the tribal confrontations rooted in history, while the exogenous violence is due to the influence of the insurgents.

The mathematical model is based on the principal-agent paradigm and captures the key cause-and-effect relationships behind work-for-security arrangements, in the context where the area of interest comprises two tribal groups.

In order to determine efficient contracts, it is key to understand the effort levels in the absence of any payments. The literature refers to this situation as a agent’s reservation utility. It is found that, depending on the affinity toward violence and the cost of exerting effort, the agents settle in either a strong equilibrium, a weak (unstable) equilibrium, or in a cycle. A strong equilibrium results if at least one of the agents has a firm decision on exerting or not an effort. Otherwise, they will find an unstable equilibrium or a cycle among the different states. This equilibrium depends not only on the attitude of each agent toward the violence (endogenous or exogenous), but also by the other agent.

In the simplest mode, the government observes the agents’ reservation utility, so its problem reduces to finding the cheapest transfers that the agents will accept; i.e., transfers that have a larger expected value than the reservation utility. These constraints are known in the literature as the agents’ participation constraints. This scenario is used to determine a baseline for the transfers offered by the principal. We found that the optimal transfer, when viewed as a function, is constant, when the agents are risk averse to money. Although this result does not provide any intuition about the most economical approach to get the agents to exert effort, the analysis of the model described scenarios in which the principal does not need to offer any inducements to obtain the cooperation of the agent(s).

The second — more realistic — scenario captures the lack of knowledge about the effort exerted by the agents. In this scenario the principal cannot play the take-it-or-leave-it game with the agents summarized in the previous paragraph. The presence of moral hazard adds more complexity, and results in an extra cost (the information rent) to the principal, because the principal has to consider incentive constraints that make the transfers sufficiently large, and thus advantageous for the agents to exert effort.
I would like to express my deep appreciation to my thesis advisor, Professor Roberto Szechtm-
man, for his support and guidance in the development of my thesis. His mentorship, insight,
and patience made this thesis possible.

I would like to thank my wife, Carmen, and my children, Federico, Javier and Carmen. I would
not be here today without your support and patience.
CHAPTER 1:
INTRODUCTION

1.1 Introduction
Conflicts between different groups where at least one group tries to influence the actions of the others have been widely studied in recent years. Situations such as humanitarian interventions [2], violence among street gangs [3], domestic terrorism [4], military conflicts, and the fight against corruption [5] were analyzed using statistical models, agent-based simulations, or analytical models. Here, the principal-agent model is used to analyze a situation comprising several stakeholders. In this model there is a reference player, the principal, who takes certain actions aimed at reducing the level of violence. The other players, the agents, decide whether to cooperate with the principal depending on the incentives offered by the principal — if their current situation improves — or if they simply do not have any other options. We present a model that takes into account the utility of the principal and the agents’ situational constraints that must be satisfied in order to obtain their cooperation. We also consider an exogenous element that possesses the ability to influence the agents to change their behavior.

This thesis is motivated by the current situation in Afghanistan, which began in November of 2001 when the Taliban were routed from Kabul, retreating toward the mountainous border region between Afghanistan and Pakistan. Since then, the United States government and its allies have attempted to stabilize Afghanistan. However, this is not an easy task due to the fragmentation of its population into myriad ethnic, linguistic, religious, and regional tribal groups. Rivalry and even armed hostilities have been historically common between and within many of these groups.

The Afghan concept of tribe is that its members are united by kinship, have a common territory, and enjoy warm social fellowship, where all are ideally equal and political organization is acephalous. Heads of tribal groups act principally as spokesmen, but have no right to make decisions that are binding on others. Hence, they are opposed to any government with a hierarchical organization, where there are rulers and ruled populations. The Pashtuns, the largest ethnic group in Afghanistan, have a complex honor-based society. Pashtunwali (the way of the Pashtuns) is the foundation of their identity. Pashtunwali’s honor-based society is governed by the concepts of chivalry (ghayrat or nang), hospitality (melmastia), gender boundaries (purdah
or namus) and council (jirga). By adhering to Pashtunwali a Pashtun possesses honor (izzat); without honor they have no rights, protection, or support of the Pashtun community [6].

Given the historic Afghan aversion to foreign forces, it appears unlikely that the coalition forces will defeat the insurgent groups with just a military force. Victory is usually a function of the struggle between local government and insurgents, but in Afghanistan, the tribes have an important role in this fight, suggesting that effectively leveraging local communities could significantly improve counterinsurgency prospects. Gaining the support of the population, especially mobilizing locals to fight insurgents, providing information on their locations and movements, and denying insurgent sanctuary, are the sine qua non of victory in counterinsurgency warfare. Such support may be spontaneous, although that is quite rare and often temporary. By tapping into tribes and other communities where resistance already exists, local defense forces can help mobilize communities in multiple areas simultaneously [7].

We analyze the situation faced by coalition forces in many small towns and villages of Afghanistan. These coalition forces are routinely exposed to attacks from insurgents, often with active participation or help from the local population. These communities tend to be motivated by self-interest and self-sufficiency, preferring to secure their own villages rather than have outsiders do it for them. That need for self-protection is a reaction that arises in populations that have experienced continuous tribal warfare.

Each tribe organizes the local forces in different ways, depending on the goal of the forces, which implements the decisions of local jirgas or shuras. ¹ The local forces often are underpaid, and their jurisdiction is limited to the territory governed by the respective shura. The local population considers it an honor to serve in the local forces. Local forces are defensive, at a village level, and are under the control of local shuras. Their main task is to provide security in the community, and they have a predisposition to engage and negotiate. If the insurgents agree not to attack coalition forces in their territory, and if they also promise not to engage in other hostile or subversive acts, community leaders might let them pass unmolested.

The coalition forces have focused on efforts to decrease the level of violence by trying to strengthen the central government institutions, the Afghan National Police (ANP) and the Afghan National Army (ANA), as bulwarks against the insurgent groups. However, local security forces are a critical complement to these national efforts [7]. The coalition forces can use the local

¹Jirga is a temporary council established to address specific issues, while a Shura is a more permanent consultative council
shuras to exercise control over the local forces, providing them with the equipment, infrastructure, and wages needed to train, arm, and organize. In return, the locals provide manpower and intelligence, but bear no economic cost. However, it has been observed that when these local forces are seen as being directly controlled by the government, they may not enjoy local support and thus not be effective, because the government is seen as corrupt and unable to provide security or basic services [7].

1.2 Motivation and Research Focus
In this thesis, we capture the key cause-and-effect relationships behind work-for-security arrangements, where the area of interest comprises two tribal groups. The principal-agent paradigm will be used to model the relationship between coalition forces (principal) and local tribes (agents), when the coalition forces may more effectively reduce the violence by contracting the security to the local forces than by doing the work themselves.

These relationships are affected by the following issues:

- The principal and the agents may not have aligned preferences and, without incentives, the agents will follow the interest of their jirga/shura. To achieve the desired alignment, the principal must provide incentives to the agents to make cooperation sufficiently rewarding.
- The agents have private information; that is, information only known by the agents. In particular, the principal may not know the value system of the agents. This issue is known as "adverse selection" [8].
- The agents may take actions that cannot be observed by the principal. For example, the principal may not know the extent of the agents’ cooperation with the insurgents, or the effort put forth by the agents in curtailing violence. This is commonly referred to as "moral hazard" [8].
- Finally, the principal and the agents share ex post the same information (the level of violence), but no third party can observe this information, resulting in a case of non-verifiability [8].

In the basic principal-agent model [9], the principal offers certain transfers depending on the observed outcome, but not on the level of effort, which is private to the agent (third bullet point...
above). The idea is that the violence outcome depends probabilistically on the effort level and, while the principal knows the nature of this dependency (e.g., the conditional density function), it does not observe the conditioning element — the effort exerted by the agent. From the agent’s perspective, exerting effort carries a cost and he can choose to stay out of the contract, in which case the agent receives his reservation utility; that is, the utility obtained in the absence of any contract. Therefore, in order to get the agent committed with a certain degree of effort, the principal offers a contract (i.e., a transfer that is dependent on the observed outcome) that: (i) provides the agent enough expected utility to meet his reservation utility; and (ii), makes the expected utility resulting from the desired level of effort at least as large as that resulting from other levels of effort.

As stated earlier, the outcome and effort level are stochastically related, so that the risk aversion of the agent is a factor to consider in the resolution of the problem. In general, the agents’ expected utility will be a decreasing function of his risk aversion, so that to encourage the agent to accept a contract, the principal must offer him payoffs/transfers that are increasingly better (viewed as an expectation taken over the relevant violence probability measure) as the agent becomes more risk averse. In the model described in this thesis, the principal requires the cooperation of the agents in order to keep the violence low. However, such cooperation involves a cost to the agents, and depends on the effort required to combat violence, the constitution of appropriate local forces, and the endogenous tendency of the agents toward violence. Thus, the principal must offer a contract that takes into account these efforts, and which results in a larger expected utility to the agents than the one they could obtain by rejecting the contract. In this thesis, the contract is individualized to each agent and there is no collusion between the agents.
1.3 Contributions of This Work

The main contribution of this thesis is the introduction of a model describing what occurs when government forces attempt to induce actions that tend to lower violence from the local tribes, in a counterinsurgency situation. Two sources of violence are considered: endogenous and exogenous. The endogenous violence is due to the tribal confrontations rooted in history, while the exogenous violence is due to the influence of the insurgents.

The motivating force behind this thesis is the situation in Afghanistan, where the fight against the insurgency requires a complex set of solutions, including non kinetic actions. In this light, this model is an attempt at capturing the impact of non kinetic actions, calibrated to obtain a desired degree of cooperation from the recipient population, taking into account their intrinsic values and the impact of a hostile external element. As a result, it is expected that this model can serve as a basis for the allocation of resources associated with non-kinetic activities.

Two possible scenarios are considered: The simplest is that of full information, where the agents do not have private information and their actions can be verified. This scenario is used to determine a baseline for the transfers offered by the principal. Knowing this baseline, the second — and more realistic — scenario captures the principal’s lack of knowledge about the actions taken by the agents. The last model provides qualitative insights about the cost to the principal due to unknown information.

1.4 Structure of the Thesis

This thesis is organized as follows. Chapter 2 provides background information on some applications of the principal-agent model and the insurgency in Afghanistan. Chapter 3 presents the model framework. Chapter 4 deals with the baseline situation in the absence of any contract. Chapter 5 analyzes the model under two perspectives: full knowledge of the principal about the private information of the agents and their actions, and partial information due to the ignorance of the agents’ outcomes. Chapter 6 provides some conclusions, and discusses potential future work in the area.
CHAPTER 2:
BACKGROUND

In this chapter, a summary is provided of the main references corresponding to the socio- political context to which this thesis applies, as well as those relevant to related agency problems.

2.1 Socio-Political Context

Afghanistan is traditionally run as a loosely knit federation with a relatively weak central government based in Kabul. No single group comprises a majority of the population, but the largest community is the Pashtun, as can be appreciated in the Afghanistan Ethno Linguistic Map in Appendix A. This ethnic group, like most others in the country, is not limited to the borders of Afghanistan, but also constitutes a major ethnic group in Pakistan. The largest single tribe of the Pashtun ethnicity, the Ghalji or Ghilzai, forms the backbone of the Taliban insurgency. The Pashtuns are fiercely independent and often view themselves, being the largest ethnicity in the country, as the rightful leaders of Afghanistan. The Hazara form the second largest group in the country. The Hazara have experienced discrimination at the hands of the Pashtun-dominated government throughout the history of Afghanistan. Hazara have been exploited as servants and laborers. As a result, they tend to be anti-government and anti-Pashtun. Due to atrocities committed against them by the Taliban, the Hazara are, by and large, opposed to the Taliban [10].

An insurgency is primarily a struggle for power between a state (or an occupying power) and one or more popularly based challengers. Since liberating Afghanistan from foreign occupation is now part of the Afghan traditions, and there is no strong government to seek the support of the tribes, Afghanistan is viewed as a good place for insurgencies to thrive. The goal of an insurgency is to mobilize human and material resources in order to form an alternative to the state – the counterstate. The counterstate may have much of the infrastructure possessed by the state itself, but this must normally be hidden, since it is illegal. As the insurgents gain confidence and power, the clandestine infrastructure may become more visible [11].

There are two approaches to fighting insurgencies, counterterrorism and counterinsurgency. Both try to weaken the capacity of the insurgents. Counterterrorism strategies focus on terrorist
networks, employing the most advanced assets to kill or capture key leaders. Counterinsurgency focuses on eliminating the medium within which insurgents live and conduct operations. The main disadvantage of counterterrorism operations is that they provide most of the material for insurgency propaganda, due to their collateral damage and civilian casualties. Counterinsurgency, by comparison, destroys the insurgents’ capacity to conduct operations by degrading local support. As security improves, the population turns its back on the insurgency and, without support, insurgents’ operations become impossible [12].

Security of the population is an imperative for counterinsurgency. The first step to ensuring security is to minimize the influence of the insurgency. The population is then mobilized, armed, and trained to protect itself. Effective security allows local, political, and administrative institutions to operate freely and commerce to flourish [11]. This is not an easy task, however, because the insurgents can undermine the willingness of the population through threats. In the last century, it was possible to verify the efficiency of so-called “night letters” (shabnamah) to intimidate the population and prevent their support of the central government [13]. The insurgents have adopted shabnamah as a well-tested, cost-effective method of instruction and intimidation. The insurgents regularly post such letters during the night, warning of the “wrath” villages will face if they cooperate with coalition forces or the Karzai regime. In Kandahar province, many police officers have quit, and after letters appeared threatening employees, two medical clinics were shut down, and many schools have burned down because they were teaching girls and women. The use of words so valuable to the Afghans, such as honor and faith, can be appreciated in the following letter [13]:

We inform those people of Marrof district that serve Americans day and night and show the places of the Mujahedeen to them or those who dishonor sincere Muslims of the country that American guards will not always be there and we can catch you any time. We know the name and place of every person; learn a lesson from who were loyal to Russians; if God wills soon you will come under the knife or bullet of Mujahedeen.

2.2 Agency Background
The coercion employed by the insurgents and the preferences of the agents can be modeled with the principal-agent paradigm. This model has been widely used in economic environments such as, for example, public-private partnerships where a public authority enters a long-term
contractual arrangement with a private supplier for the delivery of some services. In Europe, this approach has been used for some of the biggest infrastructure projects. Iossa [14] analyzed the delegation of services to the private sector that takes place in a moral hazard environment, where the efforts in infrastructure and service quality are not verifiable. Such delegation is costless when the private firm is risk neutral, but becomes costly when the agent is risk averse. When the government is risk neutral, its goal is to maximize an expected social welfare function, defined as the social benefit of the service net of its costs and the payment made to the private firm. Under complete information (when the principal knows the agent’s actions), the government offers a contract defined as a pair (effort in infrastructure quality, effort in service quality) and a reward such that the firm’s expected profit is zero and maximizes the social welfare (so-called first best outcome). Under asymmetric information, when only the agent knows its actions, the government must offer a contract that is acceptable for the firms, meeting certain incentive constraints. The optimization of social welfare, taking into account the incentive constraints, results in the second best outcome. When government transfers cannot be used to provide incentives, higher investments in infrastructure quality can be induced either by raising the fees charged to users for the service, or by increasing the length of the contract. If the government decides to increase the length of the contract, then the contract must be modified to ensure that future revenues are sufficient to cover the initial investment of the firm.

In economics, scholars have distinguished between two types of principal-agent problems: those resulting from hidden actions (moral hazard), and those resulting from hidden information (adverse selection). These problems are relatively common in third world countries, where the weak nature of political institutions makes contracting difficult. Rauchhaus [15] studied these types of problems in the context where a third party must credibly intervene when a domestic minority is persecuted or injured. At the same time, the principal must maintain the right to abstain from intervention in the event that the domestic minority is causing violence (called a commitment dilemma). In the case of humanitarian intervention, the principal offers a security guarantee (contract) to a domestic group that is at risk (the agent). In situations that entail moral hazard, the principal is unable to monitor the domestic group’s actions (provocation or restraint), and an incentive exists for the domestic minority to engage in violent or provocative behavior. Before entering a contract, the principal is not completely informed about the domestic minority’s preferences (asymmetric information), willingness to engage in dialogue (low risk), or to use violence (high risk). In this situation, the principal designs a security guarantee that takes into consideration the likelihood of each scenario. Even if the domestic
minority accepts the contract, they can engage in dialogue or violence. The most direct way for the principal to avoid moral hazard is with improved monitoring. For adverse selection, the principal should determine the motives of the domestic minorities, and screen out the risk-takers and trouble-makers.

Another example of the use of violence to provoke a humanitarian intervention can be found in the Balkans. The source of the problem is The Responsibility to Protect (R2P) international norm. The objective of this norm is to prevent and stop genocides, war crimes, ethnic cleansings and crimes against humanity. Conceptually, this is a great rule which, if used properly, would protect human rights of minorities. However, some ethnic groups in the Balkans found in this norm a way to get support from the international community. They considered engaging in the risky behavior of launching rebellions to provoke genocidal retaliation by the Serbs in the hope of achieving a humanitarian intervention. Kuperman [2] found evidence of moral hazard in several conflict zones, but the most documented cases are Bosnia and Kosovo. In both cases, genocidal violence targeted ethnic groups (Bosnia’s Muslims and Kosovo’s Albanians) after these groups pursued armed secession. Kuperman concluded that a strategy to reduce moral hazard is to declare that the international community withholds humanitarian intervention until a certain threshold of violence has been exceeded. Another strategy could be an intervention in proportion to violence, but always less than necessary to fully protect civilians or enable secessionist victory.

Leruth [5] proposes several measures to control moral hazard. He analyzed the situation in developing countries where the international community’s help for Heavily Indebted Poor Countries (HIPC’s) depends on their poverty reduction strategy, and tracking public expenditure management is critical. In this case, the Ministry of Finance (MOF) can act as the principal, providing funds to the rest of the ministries (agents) to implement a set of actions. The principal and the agents are assumed risk neutral. This model, as in any principal-agent relationship, entails hidden actions (productive effort or corruption) and hidden information (poor program design). The principal can use two control instruments, ex post audits and ante controls, and assess their value to deter cheating. With ex post audits, the auditor observes an imperfect signal on the ministry’s effort (complied, cheated) that can be used to decide to carry out the audit or not. An efficient audit reduces the agency costs (distortions and rents) associated with the second-best contract (under imperfect information, and the effort is not observable). It can be assumed that a ministry incurs a certain cost when it cheats. With ex ante controls, the MOF increases the ministry’s cost of cheating, but there is a point at which it is no longer efficient to increase
the controls. In general, the money spent on internal controls tends to be more effective than the money spent on ex post controls in developing countries. However, if the efficiency of the controls is doubtful, the ministries should be granted a rent in the form of transfer above the compensation for the effort made.

The classic Principal-Agent model only considers the relationship between one principal and one agent. But in the real world, two or more agents are usually involved, or one agent with multiple tasks. Itoh [16], [17] studied a model with multiple agents to determine when the coalition of agents can improve the principal’s benefits. In his model, each agent chooses his own effort level, which stochastically improves the outcome of the task for which he is responsible. Agents also choose the help to extend to other agents which improves the outcomes of their tasks. By selecting appropriate wage schedules, the principal can design a teamwork, in which agents are motivated to help each other. Considering only two agents, each agent simultaneously chooses a level of effort, which jointly determines the probability distribution of the outcomes of the tasks. The effort level chosen is a two-dimensional variable (own effort, helping effort). The first-best solution (under complete information) of the principal optimization problem is the effort combination which the principal wants to implement when he knows the effort chosen by each agent. In the second-best situation (under asymmetric information), the principal uses the observation of the outcomes to provide each agent with incentives to choose the desired effort levels. If an agent marginal disutility of performing an additional task is zero, the principal could secure a sufficient condition for teamwork to be optimal. Itoh determines that teamwork is optimal if each agent increases his own effort responding to an increase in help from the other agent. However, when agents’ efforts are mutually unobservable, the possibility of coalition never makes the principal better off.

Violence stemming from envy can be modeled and used as a parameter in the principal-agent paradigm. Goel [18] introduced envy into the model to narrow the gap between theory and observed characteristics of world contracts. The principal, knowing there exists envy amongst agents, studies its effect on the expected payoff. On the one hand, envy makes it easier to provide agents incentives to work hard (incentive effect). Since providing incentives to risk-averse agents is costly, this weakening of incentives lowers expected contracting costs and increases the principal’s expected payoff. On the other hand, envy reduces the expected utility of the agents, and the principal must compensate for this with higher wages (direct utility effect). Since the increase in the envy-related utility of the agent earning more is less than the decrease in the envy-related utility of the agent earning less, on average, envy reduces utility. An agent’s
utility function depends on his own wage, the envy, and the disutility function (the cost due to the agent’s effort). Each agent has a reservation utility (minimum accepted by the agents), which depends on the degree of envy, because agents’ opportunities may be affected by the degree of envy. Then, the principal’s problem is to induce agents’ actions and wage contracts to maximize his expected net payoff, subject to the agents’ individual rationality constraints (their expected utility must be greater than their reservation utility) and the agents’ incentive compatibility constraints (the expected utility with the contract is better than any outside the contract). If envious agents are able to sabotage the outcomes of other agents, the principal will have stronger incentive to compress wages and avoid the cost of such destructive actions.

There are several models dealing with violence within a principal-agent framework. Rauchhaus [15] analyzed the violence as a means to get the participation of the principal in a humanitarian intervention. Azam [19], on the contrary, focused his analysis on the way some countries use terrorism to get aid from the international community. The importance of the links amongst weak governments, aid, and terrorism was highlighted by President Bush [20] in his national security strategy: “The events of September 11, 2001, taught us that weak states, like Afghanistan, can pose as great a danger to our national interests as strong states. Poverty does not make poor people into terrorists and murderers. Yet poverty, weak institutions, and corruption can make weak states vulnerable to terrorist networks.” Azam identified in his model three players: the terrorist group that determines the number of attacks perpetrated against the donor country of the aid, the local government that exerts effort to deter these actions, and the donor that provides aid to compensate the government. Then, the donor is seeking to maximize its given income with the minimum aid and attacks costs, subject to the participation constraint of the government (the received aid must be greater than the effort cost), and the terrorist utility function (the value of the attacks must be greater than the attacks cost). The asymmetric information does not work with this model because the government involved collaborates completely by sharing intelligence. Then, under complete information, the donor does not leave any positive rent to the recipient government.

There are some situations in which several principals simultaneously and noncooperatively attempt to influence a common agent, and whose preferences for the agent’s actions usually conflict. The literature refers to this situation as a common agency. In these cases, the analysis becomes complicated because the agent knows about the schemes that have been offered by all principals at the time that he communicates with any of them. Bernheim [21] analyzed this model and considered an environment consisting of several risk-neutral principals and a single
agent. The agent may take an unobservable action that determines the probability distribution of monetary rewards received by the various principals, or may choose not to participate at all. Each principal constructs his incentive scheme in two steps: he first undoes the offers of all the other principals, and then he makes an offer. If there exists an equilibrium, each principal must select an aggregate offer that implements the equilibrium action at minimum cost. Bernheim showed that the aggregate incentive scheme induces the agent to select the equilibrium action. Peters [22] analyzed the same problem, but from another perspective. Instead of offering each principal an incentive scheme, they can construct a menu of alternatives. This menu is based on the revelation principle. Basically, the principals begin the process by designing mechanisms to establish subsequent communication. These mechanisms must be constraints to the agent. Then, the agent communicates with the principals, sending messages about his preferences. The principals respond to him by taking actions, according to what is defined in the mechanism. The agent ends the process by taking what he perceives to be the best action, and then the payoffs are realized. However, there can be some infeasible contracts. Since the agent knows what action the principal will take in response to any effort, there is no need to send more messages.
CHAPTER 3:
MODELING FRAMEWORK

3.1 Basic Situation
The situation comprises three major players: the government and coalition forces (the principal), and the two dominant tribes or clans in the area of interest (the agents A and B). Each tribal group is ruled by its shura or sheik, and behaves as a block. There exists an external player (e.g., the Taliban), who may directly or indirectly (by affecting the behavior of the tribes) induce more violence. This can occur as part of the struggle to gain the support of the tribal groups or to prevent them from cooperating with the government forces. The goal of the principal is to achieve security by mobilizing the agents, possibly by paying for their cooperation. However, the principal does not observe the cooperation directly — only the resulting violence is observable. The principal offers a contract to the tribes, at the ex ante stage, consisting of a payment amount for each violence level. Offering a generous contract would surely induce tribal cooperation but is probably unnecessarily costly, while being stingy may not get the tribes to cooperate, resulting in more violence. Thus, from the principal’s perspective the goal is to offer a menu of payments that strikes the right balance between outlays (payments to the tribes) and violence levels. This thesis studies a single-period contract and, while the contracting situation is modeled just once, the model can be generalized to multiple contracting epochs.

The presentation of the model is divided into three parts: the generation of violence, the perspective of the principal, and the stand of the agents.

3.2 Violence Generation
Let \( x = (x^{(A)}, x^{(B)}) \) be the level of violence generated by agents A and B, as measured by the number of acts of violence over a certain period of time (the duration of the contract). The element \( x^{(i)} \) is treated as a random variable with a distribution that depends on the amount of effort exerted by the agents \( e = e(e^{(A)}, e^{(B)}) \) to prevent violence, on the endogenous propensity to inter-tribal violence, \( \gamma \), and on the influence of external elements on each agent to generate violence against the principal, measured by a parameter \( \theta = (\theta^{(A)}, \theta^{(B)}) \).

To keep the model tractable, two effort levels are posed \( e^{(i)} \in \{0, 1\} \), for \( i = A, B \), with 0 meaning a low effort level, and 1 a high level of effort. The parameter \( \gamma \) can be considered
constant and known to all the stakeholders, because it can be determined based on the conflicts generated by the agent or the grievances they have suffered throughout history. It ranges from friendship ($\gamma = -1$) to enmity ($\gamma = 1$) between agents, and is reciprocal among the agents. The parameter $\theta^{(i)}$ captures the influence of external elements (e.g., the Taliban) over the agents in the generation of violent acts carried by the agents or by the external elements with tribal support; a value of $\theta^{(i)} = -1$ corresponds to animosity to the external elements, $\theta^{(i)} = 0$ reflects indifference, and $\theta^{(i)} = 1$ corresponds to friendly ties. For simplicity, it is assumed the value of $\theta^{(i)}$ is known by all parties.

While the distribution of the violence levels, $P_{e(e),e(e)}((x^{(A)}, x^{(B)})) \in d(x^{(A)}, x^{(B)}); \gamma, \theta)$, depends on $\gamma$ and $\theta$, this dependence is generally omitted to keep the notation simple and write $P_{e(e),e(e)}((x^{(A)}, x^{(A)})) \in d(x^{(A)}, x^{(B)}))$. The relationship among the variables is illustrated in Figure 3.1, which emphasizes the fact that the violence level is random, and depends on the effort level and the endogenous and exogenous violence factors.

![Figure 3.1: Dependence among key variables.](image)

The expected total level of violence, $E_{e(e),e(e)}(x^{(A)} + x^{(B)})$, tends to increase with $\theta^{(i)}$ and $\gamma$,}

16
and to decrease with $e^{(i)}$, for $i = A, B$. The effort realized by the agents is considered the private information of the agents; this information asymmetry arises when the principal can not observe and/or verify the agents’ actions, as generally occurs in situations where the principal is not totally in control of the situation.

3.3 The Principal

The principal can distinguish the source of violence, meaning that it observes the values $x^{(A)}$ and $x^{(B)}$, but only cares about the total level of violence $x^{(A)} + x^{(B)}$. The principal’s disutility due to the violence is captured by a function $S(x^{(A)} + x^{(B)})$, assumed to be strictly convex and increasing because the principal is violence averse. This reflects political considerations, where by the principal’s sensitivity to violence is non-decreasing.

The principal can make transfers to the agents to induce their effort in reducing the violence; the transfers can be in the form of wages, materials, direct payments to the tribe leader, etc. However, the effort levels $e$ are private to the agents, so that we model the transfers as real functions, $t^{(A)}(x^{(A)}, x^{(B)})$ and $t^{(B)}(x^{(A)}, x^{(B)})$, that depend just on the observable violence. Since the goal of the principal is not to punish the agents or to otherwise create a disutility for the agents, it is assumed that the transfers are non-negative. Also, since in general the principal resources are much larger than those of the agents, the transfer functions are not upper bounded. This ensures that by setting the transfers sufficiently large the principal always can induce the agents to exert effort.

The total disutility to the principal is separable in the disutility emanating from the violence and the transfers to induce cooperation from the agents, so that

$$V(x^{(A)}, x^{(B)}) = S(x^{(A)} + x^{(B)}) + t^{(A)}(x^{(A)}, x^{(B)}) + t^{(B)}(x^{(A)}, x^{(B)})$$

It is assumed the principal is risk neutral, so he only cares about the expected disutility, which is given by

$$E_{e^{(A)},e^{(B)}}[V(x^{(A)}, x^{(B)})] = E_{e^{(A)},e^{(B)}}[S(x^{(A)} + x^{(B)})] + E_{e^{(A)},e^{(B)}}[t^{(A)}(x^{(A)}, x^{(B)})] + E_{e^{(A)},e^{(B)}}[t^{(B)}(x^{(A)}, x^{(B)})]$$

(3.1)

where the expectations are taken with respect to the distribution $P_{e^{(A)},e^{(B)}}((x^{(A)}, x^{(A)}) \in d(x^{(A)}, x^{(B)}))$. The problem of the principal is to characterize the functions $t^{(A)}$ and $t^{(B)}$ that minimize his ex-
pected disutility. To make the problem non trivial, it is assumed that \( E_{0,0}[S(x^{(A)} + x^{(B)})] > E_{1,0}[S(x^{(A)} + x^{(B)})] \), \( E_{0,1}[S(x^{(A)} + x^{(B)})] > E_{1,1}[S(x^{(A)} + x^{(B)})] \), meaning that low levels of effort gives rise to more violence (or average) than high levels of effort.

When the random variable \( x^{(A)} \) has a marginal density, Figure 3.3 illustrates the effect of the agents’ effort on the marginal density of \( x^{(A)} \), showing that small values of \( x^{(A)} \) become likelier as the agents increase their violence-control efforts.

![Figure 3.2: Distribution of the violence from agent A.](image)

### 3.4 The Agents

Looking at the agents, exerting effort \( e^{(i)} \) produces a disutility \( \psi^{(i)}(e^{(i)}) \), normalized so that \( \psi^{(i)}(0) = 0 \) and \( \psi^{(i)}(1) = \psi^{(i)} \). The disutility brought upon by exerting effort captures compensation to clan leaders within a tribe, material for enforcement, etc. When the agent exerts an effort, it knows the probability distribution of the violence but not the resulting violence that results from the effort which, as discussed earlier, is presumed random.

The agents’ utility is determined by the transfer from the principal, the utility of violence, and the effort exerted to prevent violence, if any. Since the violence has an exogenous and endogenous component (as measured by \( \theta^{(i)} \) and \( \gamma \), respectively), we treat the agents’ utility as separable in the transfer, effort, and the benefits they can obtain from violence,

\[
U^{(i)}(x^{(A)}, x^{(B)}; \gamma, \theta^{(i)}) = u^{(i)}(t^{(i)}(x^{(A)}, x^{(B)})) - \psi^{(i)}(e^{(i)}) + h^{(i)}(x^{(i)}, x^{(j)}, \gamma) + g^{(i)}(x^{(i)}, \theta^{(i)}),
\]

(3.2)

where the function \( u \) is the agents’ utility of the transfer (presumed strictly increasing and concave), the function \( h^{(i)} \) captures the endogenous animosity between the agents, and \( g \) represents
the exogenous animosity caused by the insurgents. In other words, \( h^{(i)}(\cdot) + g^{(i)}(\cdot) \) reflects the interest of agent \( i \) in the violence.

Agent \( i \) for \( i \in \{A, B\} \), maximizes its expected utility, given by

\[
E_{e(A),e(B)}[U^{(i)}(x^{(A)}, x^{(B)}; \gamma, \theta^{(i)})] = E_{e(A),e(B)}[u^{(i)}(t^{(i)}(x^{(A)}, x^{(B)}))] - \psi^{(i)}(e^{(i)}) + E_{e(A),e(B)}[h^{(i)}(x^{(i)}, x^{(j)}, \gamma)] + E_{e(i)}[g^{(i)}(x^{(i)}, \theta^{(i)})]. \tag{3.3}
\]

Since \( \gamma \) and \( \theta \) are treated as constants, their effect is manifested by their impact on the functions \( h^{(i)} \) and \( g^{(i)} \), respectively.

An agent accepts the contract offered by the principal if its expected utility under the contract is at least the reservation utility. The reservation utility of an agent is the expected utility attainable by the agent before entering any contract with the principal, and is treated next.

To simplify the notation, for the remainder of this work we omit the arguments from \( U^{(i)}, t^{(i)}, h^{(i)}, \) and \( g^{(i)} \) whenever their meaning remains clear.
THIS PAGE INTENTIONALLY LEFT BLANK
CHAPTER 4:  
RESERVATION UTILITY

In the absence of a contract by the principal, the utility of each agent hinges on the effort applied to lower the violence and the utility of violence (cf., Eq. 3.2), with the latter depending on the violence generated by each agent. This suggests that the effort level chosen by each agent follows certain game dynamics. More precisely, at stage $t$ agent $A$ decides whether to exert effort depending on the effort of agent $B$ at time $t - 1$; agent $B$ simultaneously takes a similar decision by observing the effort level of agent $A$ at time $t - 1$. The decision taken by each agent at time $t$ is a best response, meaning that an agent maximizes the utility that would be obtained if the other agent did not change its level of effort.

These dynamics result in three possible outcomes: (i) a strong equilibrium, whereby any initial effort level eventually results in the same effort level decisions; (ii), a weak equilibrium, that arises from best-responding over subset of initial effort levels (that is, there exists at least two initial effort levels that lead to two different outcomes); and (iii), a cycle, meaning that the sequence of best responses follows a cycle over the effort level combinations, regardless of the initial conditions. In cases (i) and (ii), the dynamics result in an equilibrium, and each agent derives certain utility. In case (iii), each agent obtains an average utility, taken over the utilities observed during a cycle. With this in mind, we define the reservation utility as the average utility obtained by an agent that best-responds according to the above rules, in the absence of any transfer by the principal.

As it turns out, the reservation utility depends on whether the following inequalities are satisfied for agent $A$,

$$-\psi(A) + E_{11}[h(A)] + E_1[g(A)] \geq E_{01}[h(A)] + E_0[g(A)]$$ (4.1)

$$-\psi(A) + E_{10}[h(A)] + E_1[g(A)] \geq E_{00}[h(A)] + E_0[g(A)]$$ (4.2)

and for agent $B$,

$$-\psi(B) + E_{11}[h(B)] + E_1[g(B)] \geq E_{10}[h(B)] + E_0[g(B)]$$ (4.3)

$$-\psi(B) + E_{01}[h(B)] + E_1[g(B)] \geq E_{00}[h(B)] + E_0[g(B)]$$ (4.4)
These inequalities are the key driving element behind the best response dynamics. For instance, if (4.1) holds, then agent A is better off exerting effort when agent B does too.

To simplify the notation, we let $c(a, b)$ indicate the number of inequalities satisfied by agents A and B, with $a, b \in \{0, 1, 2\}$. It will be seen that the effort level dynamics lead to a strong equilibrium if $a \in \{0, 2\}$ or $b \in \{0, 2\}$; that is, if at least one agent has a firm decision about exerting effort or not. Otherwise, if $c(1, 1)$, the effort level dynamics result in a weak equilibrium or in a cycle. The argument proceeds as follows.

- **Case $c(2, 2)$**: Both agents exert an effort regardless of the action of the other, and $(e^{(A)}, e^{(B)}) = e(1, 1)$ is a strong equilibrium.

- **Case $c(2, 1)$**: The inequalities (4.1) and (4.2) are satisfied, but only one of (4.3) and (4.4) is satisfied. We find a different equilibrium depending on the priorities of the agents. All the cases can be classified into two big groups, those that meet the inequality (4.3) and those that do not.

  - Only inequality (4.3) holds for agent B. Agent A exerts effort regardless of B’s effort level. It can be seen that agent B is better off exerting effort when agent A does too, so that the equilibrium is $e(1, 1)$.

  - Only inequality (4.4) holds for agent B. As above, agent A always exerts effort, but agent B is better off not applying effort, resulting in $e(1, 0)$ being a strong equilibrium.

- **Case $c(2, 0)$**: The agent A has a firm decision to exert an effort, and the agent B not to apply it, regardless of the other agent’s actions. Hence $e(1, 0)$ is a strong equilibrium.

- **Case $c(0, 0)$**: Both agents are better off not applying effort, so that $e(0, 0)$ is a strong equilibrium.

- **Case $c(1, 0)$**: Agent B is better off not applying effort, regardless of the actions of agent A.

  - Only inequality (4.1) holds for agent A: Since (4.2) does not hold, agent A is better off not applying effort in response to a lack of of effort from agent B. Therefore, $e(0, 0)$ is a strong equilibrium.
– Only inequality (4.2) holds for agent $A$: Agent $A$ is better off exerting effort in response to a lack of effort from agent $B$. Therefore, $e(1,0)$ is a strong equilibrium.

• (1,1): Only one of the inequalities in (4.1–4.2) and in (4.3–4.4) are satisfied. There are four possible scenarios.

– Only inequalities (4.1) and (4.3) hold: Agent $A$ responds to $e^{(B)} = 1$ with $e^{(A)} = 1$ and to $e^{(B)} = 0$ with $e^{(A)} = 0$. Likewise, agent $B$ responds to $e^{(A)} = 1$ with $e^{(B)} = 1$ and to $e^{(A)} = 0$ with $e^{(B)} = 0$. Hence, if the initial set of actions is $e(1,1)$, the equilibrium remains at $e(1,1)$, and if the initial set of actions is $e(0,0)$, the equilibrium remains at $e(0,0)$. On the other hand, if initial actions are $e(1,0)$, the best responses lead to $e(0,1)$, which results in $e(1,0)$, so we have a cycle.

– Only inequalities (4.2) and (4.4) hold: In this case agent $A$ responds to $e^{(B)} = 1$ with $e^{(A)} = 0$ and to $e^{(B)} = 0$ with $e^{(A)} = 1$. Likewise, agent $B$ responds to $e^{(A)} = 1$ with $e^{(B)} = 0$ and to $e^{(A)} = 0$ with $e^{(B)} = 1$. This situation leads to a cycle $e(0,0) \rightarrow e(1,1) \rightarrow e(0,0)$, or weak equilibriums $e(1,0) \rightarrow e(1,0)$ or $e(0,1) \rightarrow e(0,1)$

– Only inequalities (4.1) and (4.4) hold: Agent $A$ responds to $e^{(B)} = 1$ with $e^{(A)} = 1$ and to $e^{(B)} = 0$ with $e^{(A)} = 0$. Agent $B$ responds to $e^{(A)} = 1$ with $e^{(B)} = 0$ and to $e^{(A)} = 0$ with $e^{(B)} = 1$. A cycle must occur in this case: $e(1,1) \rightarrow e(1,0) \rightarrow e(0,1) \rightarrow e(1,1)$, or $e(0,0) \rightarrow e(0,1) \rightarrow e(1,1) \rightarrow e(1,0) \rightarrow e(0,0)$.

– Only inequalities (4.2) and (4.3) hold: Agent $A$ responds to $e^{(B)} = 1$ with $e^{(A)} = 0$ and to $e^{(B)} = 0$ with $e^{(A)} = 1$. Agent $B$ responds to $e^{(A)} = 1$ with $e^{(B)} = 1$ and to $e^{(A)} = 0$ with $e^{(B)} = 0$. A cycle must occur in this case: $e(1,1) \rightarrow e(0,1) \rightarrow e(0,0) \rightarrow e(1,0) \rightarrow e(1,1)$, or $e(0,0) \rightarrow e(1,0) \rightarrow e(1,1) \rightarrow e(0,1) \rightarrow e(0,0)$.

If the equilibrium is $e(1,1)$, the principal is in an optimal situation because the agent exerts the highest effort for free, in the absence of any transfer, and the principal does not need to propose a contract. In the remaining situations, the principal has to offer an incentive to at least one agent in order to induce an equilibrium $e(1,1)$. Whether such inducement is worthwhile depends on the utility function of the principal, and on the expected cost of the transfer needed to induce effort. These trade-offs are studied in the next chapter, first in the context of complete information, and later in the incomplete information scenario.
In this chapter, we study the contracts that are needed to induce high level of effort from both agents in two situations, perfect and imperfect information.

5.1 The Case of Full Information

This section treats the situation when the principal can observe the effort exerted by the agents, which has the advantage of being easily tractable and serves as a benchmark for determining the (so-called) information rent extracted by the agents when there is incomplete information. In this scenario, the principal wishes to determine the least expensive transfers (on average over the levels of violence) that ensure the agents exert a high level of effort. It was seen earlier, in Chapter 4, that a high level of effort can result when either (4.1, 4.2, and 4.3) or (4.1, 4.3, and 4.4) hold in the absence of transfers. Therefore, for a given baseline situation, there generally exists more than one way leading to $e(1, 1)$ situation. Thus, the goal of this chapter is to study these possible contracts, and to gain insights about the most efficient modus operandi that guarantees a high level of effort from the agents. Throughout, it is assumed that the principal knows the value of $\theta$ before offering the menu of contracts.

The timeline of the contract, shown in Figure 5.1, is as follows: (i) the agents and principal discover the value of $\theta$, and $\gamma$ is common knowledge; (ii) the principal offers to each agent a menu of contracts based on the level of violence that emanates from its community; (iii) the agents accept the contract if it meets the minimum participation requirements (cf., Eq. (5.1)); (iv) the agents exert a positive effort, and the realized (random) output is $(x(A), x(B))$; and lastly, (v) the principal executes the contract, (i.e., realizes the transfer).

![Figure 5.1: Time line of the contractual game.](image)

The principal must offer the agents transfers that result in a utility level that is at least their
reservation utility in order to get them to participate in the violence lowering efforts. These constraints are known in the literature as the agents’ participation constraints, and can be written as

\[ E_{1,1}[U^{(i)}] \geq \bar{U}^{(i)}, \ i = A, B, \]  

(5.1)

where \( \bar{U}^{(i)} \) is the reservation utility of the agent \( i \). As the principal has full information, he only checks the predisposition of the agents to exert or not an effort.

As stated earlier in Chapter 3, the principal never punishes the agents, so that

\[ t^{(i)} \geq 0, \ i = A, B. \]  

(5.2)

Hence, if the principal wants to induce effort in each agent, the optimization problem becomes

\[
P_1 : \min_{t^{(A)}, t^{(B)}} E_{1,1} \left[ V(x^{(A)}, x^{(B)}) \right] \text{ subject to (5.1) and (5.2).} \]  

(5.3)

Clearly, the feasible set is non empty, as the transfers can be made arbitrarily large in order to satisfy the participation constraints (5.1). Moreover, if \( E_{1,1}[h^{(i)}] + E_{1}[g^{(i)}] \geq \psi^{(i)} + \bar{U}^{(i)} \) then setting \( t^{(i)} = 0 \) is optimal because any value of \( t^{(i)} \) lower than 0 would be unfeasible. Otherwise, if \( E_{1,1}[h^{(i)}] + E_{1}[g^{(i)}] < \psi^{(i)} + \bar{U}^{(i)} \), then (5.1) does not hold in the absence of any transfers, so that a feasible solution is given by setting a constant transfer,

\[
\tilde{t}^{(i)} = u^{(i)^{-1}} \left[ \psi^{(i)} + \bar{U}^{(i)} - E_{1,1}[h^{(i)} + g^{(i)}] \right].
\]  

(5.4)

This solution satisfies (5.2) and, computing \( E[u^{(i)}(\tilde{t}^{(i)})] \), can be seen to lead to \( E_{1,1}[U^{(i)}] = \bar{U}^{(i)} \), so that (5.1) holds. That \( \tilde{t}^{(i)} \) is optimal follows from the concavity of \( u^{(i)} \). Indeed, suppose that \( \check{t}^{(i)} \) is another optimal solution such that \( \check{t}^{(i)} \neq \tilde{t}^{(i)} \) over a set \( \Gamma \), with \( P((x^{(A)}, x^{(B)}) \in \Gamma) > 0 \). If \( \check{t}^{(i)} \) is optimal then the participation constraint (5.1) must be tight, for otherwise it would be possible to find cheaper transfer functions for which the participation constraint is satisfied. Thus, it must be that

\[
E_{1,1}[u^{(i)}(\check{t}^{(i)}))I(x^{(A)}, x^{(B)}) \in \Gamma] = E_{1,1}[u^{(i)}(\tilde{t}^{(i)}))I(x^{(A)}, x^{(B)}) \in \Gamma],
\]

26
where $I(\cdot)$ is the indicator function.

However, $\hat{t}^{(i)}$ constant means that

$$u^{(i)}(E_{1,1}[\hat{t}^{(i)} I(x^{(A)}, x^{(B)}) \in \Gamma]) = E_{1,1}[u^{(i)}(\hat{t}^{(i)}) I(x^{(A)}, x^{(B)}) \in \Gamma],$$

and Jensen’s inequality results in

$$E_{1,1}[u^{(i)}(\tilde{t}^{(i)}) I(x^{(A)}, x^{(B)}) \in \Gamma] \leq u^{(i)}(E_{1,1}[\tilde{t}^{(i)} I(x^{(A)}, x^{(B)}) \in \Gamma]),$$

so that

$$u^{(i)}(E_{1,1}[\hat{t}^{(i)} I(x^{(A)}, x^{(B)}) \in \Gamma]) \leq u^{(i)}(E_{1,1}[\tilde{t}^{(i)} I(x^{(A)}, x^{(B)}) \in \Gamma]).$$

Since $u^{(i)}$ is monotone increasing, applying $u^{(i)^{-1}}$ preserves the inequality, resulting in

$$E_{1,1}[\hat{t}^{(i)} I(x^{(A)}, x^{(B)}) \in \Gamma] \leq E_{1,1}[\tilde{t}^{(i)} I(x^{(A)}, x^{(B)}) \in \Gamma],$$

meaning that the transfer $\hat{t}^{(i)}$ is cheaper than $\tilde{t}^{(i)}$ for the principal, and thus optimal.

This is the transfer made by the principal to the agent $i$, and is referred to in the literature as the first-best transfer $\hat{t}^{(i)}_{FB}$.

Hence, the optimal transfer, when viewed as a function, is constant, as would be expected given the agents’ risk aversion. Unfortunately, this result does not provide any intuition about the most economical approach to get the agents to exert effort. There are several possible scenarios, depending on the reservation utility, as discussed in Chapter 4:

- If a high level of effort is a strong equilibrium for both agents in the absence of any transfers, then the principal does not offer any inducements to obtain the cooperation of the agents. This is a best-case scenario for the principal, where both agents have a self-interest in keeping violence low. In the Afghan context, the principal can leave to the shura to manage the local forces without any external interference.

- If $e(1, 0)$ is a strong equilibrium, meaning that, absent any transfers, agent $A$ exerts effort and agent $B$ does not, from Chapter 4 it is known that one the following scenarios must hold: (i), inequalities (4.1), (4.2), and (4.4) are satisfied; or (ii), when only inequality (4.2) holds. In case (i), agent $A$ cooperates always and the participation constraint (5.1)
holds for \( A \), so in order to get to \( e(1,1) \) the principal needs to pay agent \( B \) just enough so that (5.1) holds for \( B \) ((4.3) is the same as (5.1), and can be omitted). In case (ii), the principal can pay to meet (4.1) and (4.3). (Paying to meet (4.1), (4.3), and (4.4) is at least as expensive.)

- If \( e(1,1) \), from Chapter 4 is known that a cycle or weak equilibrium results. In this case, the principal has several options, depending on the situation in the absence of any transfers: (i), where inequalities (4.1) and (4.3) hold; (ii), where inequalities (4.2) and (4.4) hold; case (iii), when (4.1) and (4.4) hold; and (iv), when (4.2) and (4.3) are true.

In case (i), the principal can induce \( e(1,1) \) by paying to satisfy (4.2) or (4.4), whichever is cheapest. The same is true in case (ii), where the principal can pay to meet the cheapest of (4.1) and (4.3) to induce \( e(1,1) \). In case (iii), the principal only needs to pay agent \( B \) so that (4.3) holds. Likewise, in scenario (iv), the principal only needs to pay agent \( A \) so that (4.1) becomes valid.

- The last, most complex scenario, is when \( e(0,0) \) is a strong equilibrium. The study of the minimum transfers needed to achieve a high level of effort follows along the line of the preceding cases, and is omitted.

Keep in mind, the principal can pay to meet the inequalities (4.1, 4.2, and 4.3) or (4.1, 4.3, and 4.4) to shift the effort levels from \( e(0,0) \) to a strong equilibrium \( e(1,1) \), but this does not mean that the constraint (5.1) is satisfied.

In all the above scenarios, the shape of the transfer function is the same — constant — but the transfer amount depends on the inequalities (4.1) – (4.4) of Chapter 4. Throughout, as the principal progressively increases the transfers to both agents, the effort levels can shift from \( e(0,0) \) strong to a weak equilibrium. This is illustrated in Figure 5.2-a, where the agents’ utilities appear as a function of the transfers, when the none of the inequalities (4.1 – 4.4) hold, and in Figure 5.2-b, where agent \( A \) meets one of the inequalities, with the gray area representing a weak equilibrium.

### 5.2 Problem With Partial Information

In the full information situation, the principal knows the private information of the agents and can verify the exerted effort, so that the transfers are the smallest that meet the participation constraints. However, the agents may be able to hide the true effort to the principal and to the
other agent — a problem known in the literature as moral hazard. We treat this situation in this section and, as in the case of full information, assumed that the realized level of violence is common knowledge and that the value of $\theta$ is known to all parties.

The timeline of the contract, shown in Figure 5.3, is as follows: (i) the principal knows the tendency to violence of each agent ($\theta$ and $\gamma$), but is unable to verify whether the agent is making an effort or not; (ii) the principal offers to each agent a menu of contracts based on the expected level of violence emanating from the community; (iii) given $\theta$ and $\gamma$, the agents independently decide to exert a positive effort or not; (iv) the principal cannot verify the effort exerted, but can do some inference by observing the outcomes; and lastly (v), the principal executes the contract.

The principal wishes to minimize the expected disutility needed to induce a high level of effort, when the effort is not observed. Hence, the transfers must be sufficiently large to make it advantageous for the agents to exert effort. More precisely, the transfers must satisfy the following constraints:

- **Incentive constraints:** To prevent the agents from being better off by exerting no effort, the transfers must ensure that the agents’ expected utilities are higher when they produce high effort, regardless of the behavior of the other agent:

\[
E_{1,1}[u^{(A)}(t^{(A)})] - \psi^{(A)}_1 + E_{1,1}[h^{(A)} + g^{(A)}] \geq E_{0,1}[h^{(A)} + g^{(A)}] + E_{0,1}[u^{(A)}(t^{(A)})] \quad (5.5)
\]
\[E_{1,0}[u^{(A)}(t^{(A)})] - \psi^{(A)}_1 + E_{1,0}[h^{(A)} + g^{(A)}] \geq E_{0,0}[h^{(A)} + g^{(A)}] + E_{0,0}[u^{(A)}(t^{(A)})] \quad (5.6)\]

The corresponding constraints for agent \(B\) are:

\[E_{1,1}[u^{(B)}(t^{(B)})] - \psi^{(B)}_1 + E_{1,1}[h^{(B)} + g^{(B)}] \geq E_{1,0}[h^{(B)} + g^{(B)}] + E_{0,0}[u^{(B)}(t^{(B)})] \quad (5.7)\]

\[E_{0,1}[u^{(B)}(t^{(B)})] - \psi^{(B)}_1 + E_{0,1}[h^{(B)} + g^{(B)}] \geq E_{0,0}[h^{(B)} + g^{(B)}] + E_{0,0}[u^{(B)}(t^{(B)})] \quad (5.8)\]

- Participation constraints: The agents may not participate in the contract with the principal, in which case their reservation utility is \(\bar{U}^{(i)}\). Therefore, to induce cooperation, the principal has to satisfy the constraints

\[E_{1,1}[u^{(A)}(t^{(A)})] - \psi^{(A)}_1 + E_{1,1}[h^{(A)} + g^{(A)}] \geq \bar{U}^{(A)}. \quad (5.9)\]

For the agent \(B\), it is required

\[E_{1,1}[u^{(B)}(t^{(B)})] - \psi^{(B)}_1 + E_{1,1}[h^{(B)} + g^{(B)}] \geq \bar{U}^{(B)}. \quad (5.10)\]

It is not necessary to induce participation of an agent when the other is not exerting an effort. This can be rationalized as follows: First, the principal finds a transfer function that meets the incentive constraints, which guarantee \(e(1,1)\) only if the agents agree to participate. Last, the principal may have to increase the transfer to make it worthy for the agents to participate.

As in the full information case, we assume that the principal does not punish the agents, so the transfers cannot be negative:

\[t^{(i)} \geq 0, \text{ for } i \in \{A, B\}. \quad (5.11)\]

Therefore, the optimal transfers needed induce cooperation from the agents are a solution of

\[\textbf{P}_2 : \min_{t^{(A)}, t^{(B)}} E_{1,1}[V(x^{(A)}, x^{(B)})], \text{ s.t. } (5.6) - -(5.11).\]

The incentive constraints prevent the principal with partial information from achieving a better solution than with full information. The difference between the first best solution with full information and the second best solution with partial information is called the information rent, and reflects the cost to the principal of the moral hazard.
We observe, from the analytical standpoint, that the transfers that satisfy incentive and participation constraints form a closed and convex subset with a non empty interior, since (recall) the agents’ utility of money $u^{(i)}$ is concave. Just the same, the objective function is linear in the transfers to the agents. Therefore, the standard theory ([23]) suggests that the Karush-Kuhn-Tucker conditions are sufficient for optimality. This allows the analysis of some aspects of the problem treating each agent independently.

Without loss of generality, we assume that the level of violence $x = (x^{(A)}, x^{(B)})$ has a density $f^{ij}(x)$ for effort level $e(i, j)$. Also, $u'(t^A)$ is the partial derivative of the transfer’s utility function with respect $t^A(x)$. Focusing on the agent $A$, the first order optimality conditions for agent $A$ are

$$f^{11} - u''(A)(t^A)[\mu_1(f^{11} - f^{01}) + \mu_2(f^{10} - f^{00}) + \mu_3 f^{11}] - \mu_4 = 0$$  \hspace{1cm} (5.12)

where $\mu_1, \mu_2, \mu_3,$ and $\mu_4$ are the Lagrange multipliers of the constraints (5.6), (5.5), (5.9), and (5.11), respectively, and $\mu_i \geq 0$, for $i = 1, \ldots, 4$.

If at least one constraint (5.6), (5.5), or (5.9) is violated with $t^{(A)} = 0$, then it must be $t^{(A)} > 0$, so that $\mu_4 = 0$. In this case, from Eq. (5.12) we get

$$t^{(A)}(x) = (u''(A))^{-1} \left( \frac{f^{11}(x)}{\mu_1(f^{11}(x) - f^{01}(x)) + \mu_2(f^{10}(x) - f^{00}(x)) + \mu_3 f^{11}(x)} \right)$$  \hspace{1cm} (5.13)

where $(u''(A))^{-1}$ is the inverse function of $u'(A)$, which exists because we assume $u'(A)$ is strictly concave.
It follows that if (5.9) is the only binding constraint then \( t_{SB}^{(A)}(x) = t_{FB}^{(A)}(x) = (u')^{-1}(1/\mu_3) \), so that the second best transfers are constant, and equal to the first best transfers. In other words, if the incentive constraints are not tight at optimality, the partial information problem reduces to the complete information problem.

If the participation constraint is not binding at optimality, then \( \mu_3 = 0 \), and we conclude that \( t^{(A)}(x) = 0 \) whenever the argument inside brackets in (5.13) is negative or larger than \( u^{(A)}(0) \). Hence, agent \( A \) receives no transfers whenever \( f^{01}(x) \geq f^{11}(x) \) and \( f^{00}(x) \geq f^{10}(x) \). This can be rationalized as follows: The principal makes no transfers to agent \( A \) when there is more probabilistic evidence of \( e^{(A)} = 0 \) than of \( e^{(A)} = 1 \), regardless of the value of \( e^{(B)} \). Likewise, in regions where \( f^{01}(x) \approx f^{11}(x) \) and \( f^{00}(x) \approx f^{10}(x) \), the transfers to agent \( A \) are relatively smaller than in regions where \( f^{01}(x) \ll f^{11}(x) \) and \( f^{00}(x) \ll f^{10}(x) \) (for similar values of \( f^{11}(x) \)); i.e., where there is relatively higher evidence of agent’s \( A \) effort, which is what the principal is paying for (for similar costs \( f^{11}(x) \)). The intuition behind this is that, ceteris paribus in terms of probabilistic evidence of effort, the principal wishes to minimize \( \int f^{(11)}(x)t^{(A)}(x)dx \), which becomes larger when the integrands are jointly large.

The analysis of the situation where the participation constraint (5.9) is binding at optimality is more complex because now \( t^{(A)}(x) \) may be positive even if there is more evidence of low effort. The complete analysis of this case will be treated elsewhere.

As in the full information case, there may exist several ways for the principal to induce a high level of effort from the agents, depending on the equilibrium situation in the absence of transfers. For example, if only (4.2) and (4.4) are met, the principal has the option to pay agent \( A \) so that (5.6) is satisfied, or to pay agent \( B \) so that (5.8) holds. Since paying so that (5.6) or (5.8) hold results in \( e = e(1, 1) \), clearly the principal is better off by choosing the meet the cheapest. More precisely, the principal compares \( \int t^{(A)}(x)f^{11}(x) \) with \( \int t^{(B)}(x)f^{11}(x) \), where \( t^{(A)} \) comes from (5.13) and \( t^{(B)} \) comes from the analogue equation,

32
\[ t^{(B)}(x) = (u^{(B)})^{-1} \left( \frac{f^{11}(x)}{\lambda_1(f^{11}(x) - f^{10}(x)) + \lambda_2(f^{01}(x) - f^{00}(x)) + \lambda_3 f^{11}(x)} \right). \] 

(5.14)

where \( \lambda_i \geq 0, i \in \{1, ..., 3\} \) are the Lagrange multipliers corresponding to (5.7), (5.8), and (5.10), respectively.

Unfortunately, it appears difficult to obtain further insights unless we make stronger assumptions as, for example, stating the agents’ utilities. For other baseline equilibrium situations, the analysis is similar.
CHAPTER 6:
CONCLUSIONS AND EXTENSIONS

6.1 Motivation and Setting
It is now widely accepted that a successful conclusion to the war in Afghanistan necessitates an array of solutions that includes non kinetic actions, especially those in which the local population plays a key role. In particular, there is growing interest in having the local population take control of regional security, providing protection from external threats as well as internal sources of violence. Thus motivated, in this thesis, a new mathematical model dealing with incentives in a counterinsurgency situation was presented and analyzed. More precisely, this model captures the situation in which the coalition forces pay the locals to help reduce violence, and where the only observable are the number of violent acts. Knowing that more cooperation makes violence less likely, the problem for the coalition forces is to design a menu of payments, depending on the observed level of violence over some predetermined period of time, that makes it appealing for the local population to exert effort in reducing the violence. The coalition forces must balance the monetary costs of the incentives against the costs derived from violent acts, while the local tribes trade the cost of exerting effort to reduce violence against the potential benefits of carrying out violent acts, knowing that their level of cooperation is only indirectly (in a probabilistic sense) observed.

The mathematical model is based on the principal-agent paradigm and captures the key cause-and-effect relationships behind work-for-security arrangements, in the context where the area of interest comprises two tribal groups. We consider two sources of violence, endogenous and exogenous. The former is due to historical issues among the tribes, while the later can be ascribed to an external player. The main contribution of this thesis is the introduction of a model describing what occurs when coalition forces attempt to induce actions from the local tribes that result in less violence and, to the best of our knowledge, is the first attempt to frame this kind of situation within an agency framework.

6.2 Analysis
In order to determine efficient contracts, it is key to understand the effort levels in the absence of any payments. It was found in this thesis that, depending on the affinity towards violence and the cost of exerting effort, the agents settle into either a strong equilibrium, a weak (unstable)
equilibrium, or into a cycle. A strong equilibrium results if at least one of the agents has made a firm decision about whether or not to exert effort. Otherwise, they will find an unstable equilibrium or a cycle between the different states. This equilibrium depends not only on the attitude of each agent toward the violence (whether endogenous or exogenous), but also on that of the other agent.

In the simplest case, the government observes the level of effort exerted by the agents, so its problem reduces to finding the cheapest transfers that the agents will accept; i.e., transfers that have a larger expected value than the reservation utility. This scenario is used to determine a baseline for the transfers offered by the principal. We found that the optimal transfer, when viewed as a function, is constant, as would be expected given the agents’ risk aversion. Although this result does not provide any insights about the most economical approach in motivating the agents to exert effort, the analysis of the model described scenarios in which the principal does not need to offer any inducements to obtain the cooperation of the agent(s).

The second — and more realistic — scenario captures the lack of knowledge about the effort exerted by the agents. In this scenario, the principal cannot play the take-it-or-leave-it game with the agents summarized in the previous paragraph. The presence of moral hazard adds more complexity, and results in an extra cost (the information rent) to the principal, which is characterized under various scenarios.

6.3 Extensions
The model defined in this thesis forms a basic framework that can be extended to other situations to provide a better understanding in a counterinsurgency environment. In this section, we discuss some of these extensions briefly.

6.3.1 Many Principals and Many Agents
The model established in Chapter 4 considered one principal with two agents, and the presence of the insurgents is materialized with the parameter $\theta$. However, the problem can be expanded, rather analyzing the case where several principals compete for the favors of multiple agents.

6.3.2 Multiple Periods
The contract studied in this thesis applies over a single period, but long wars, as in Afghanistan, require multi-period models with dynamic information structures. The idea is to combine the principal-agent model with the structure of a Markov decision process. The level of violence
changes from period to period, with transition probabilities depending on the efforts realized by the agents, and the principal makes transfers to the agents based on the state transitions observed. The mathematical analysis of these models can be made more tractable if we assume that both the principal and the agents have perfect knowledge about the preferences of the other, and of the transition probabilities in the Markov process ([24]).

6.3.3 Multiple Levels of Effort
The agents in the proposed model can exert two levels of effort, high and low. It would be beneficial to expand the model so that the agents can apply effort over a range of levels.
THIS PAGE INTENTIONALLY LEFT BLANK
LIST OF REFERENCES


APPENDIX

[Afghanistan Map]

Figure A.1: Afghanistan Ethno Linguistic Map from the Institute for the Study of War. [1]
Initial Distribution List

1. Defense Technical Information Center  
   Ft. Belvoir, Virginia

2. Dudly Knox Library  
   Naval Postgraduate School  
   Monterey, California

3. Professor Roberto Szechtman  
   Department of Operations Research  
   Naval Postgraduate School  
   Monterey, California

4. Professor Michael Atkinson  
   Department of Operations Research  
   Naval Postgraduate School  
   Monterey, California