National Responses to Transnational Terrorism: 
Intelligence and Counterterrorism Provision*

Thomas Jensen†

March 21, 2014

Abstract

Intelligence about transnational terrorism is generally gathered by national agencies. I set up and analyze a game theoretic model to study the implications of national intelligence gathering for the provision of domestic (defensive) counterterrorism when two countries are facing a common transnational terrorist threat. It is shown that, relative to a benchmark case where all intelligence is known by both countries, national intelligence gathering often leads to increased inefficiencies in counterterrorism provision. By extending the main model with a communication stage, I also explore whether the differences in information that may follow from national intelligence gathering will be overcome by intelligence sharing. If verifiable sharing is a viable option for each country, then full credible intelligence sharing can happen in equilibrium. On the other hand, if only cheap talk communication is possible then it cannot.

Keywords: Transnational terrorism, counterterrorism, intelligence, intelligence sharing

---

*I gratefully acknowledge funding from The Danish Council for Independent Research | Social Sciences (grant number 09-066752).

†Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K, Denmark. Email: Thomas.Jensen@econ.ku.dk.
1 Introduction

Counterterrorism policy decisions are generally made in an environment of uncertainty. Intelligence is rarely complete, so policy makers face difficult choices when deciding whether the benefits of particular policies outweigh the costs. When the terrorist threat is transnational, such as the threat from Al-Qaeda and related groups in recent years, a new dimension of uncertainty is added. Since intelligence is generally gathered by national agencies (e.g., Walsh 2009), the authorities in one country may well be uncertain not only about the terrorists’ capabilities, but also about the intelligence collected by other countries. And since counterterrorism policies are decided on the national level, this type of uncertainty has potentially large implications for the provision of counterterrorism. In this paper I explore these implications by game theoretic modeling.

A substantial and important literature has used game theory to study counterterrorism provision when several countries are facing a transnational terrorist threat (Sandler and Lapan 1988, Rosendorff and Sandler 2004, Arce and Sandler 2005, Sandler and Siquiera 2006, and others). This literature demonstrates that counterterrorism policy-making is a strategic situation between potentially targeted countries because of transnational externalities. A central result is that defensive counterterrorism is typically oversupplied because of the negative externalities due to transference of terrorist attacks when one country’s policies make it harder to successfully attack there.¹

While the consequences of national policy-making in the face of transnational terrorism have been extensively studied, the implications of national intelligence gathering have been left largely unexplored. In the existing literature it is generally assumed that all countries have the same information.² The importance of studying the consequences of national intelligence gathering is clearly illustrated by the limited formal intelligence cooperation in the European Union. Despite a high and increasing level of integration in important areas and many common security interests, intelligence collection is still primarily done by the individual nation states and there is strong resistance to ideas of federal intelligence structures (Walsh 2006, 2009, Aldrich 2009). Thus, the barriers to formal collaboration on intelligence are apparently hard to overcome, so differences in intelligence between countries is likely to remain a common issue. This makes it important to study carefully its implications for counterterrorism policies, especially given the huge amount of resources spent in this area.

In this paper I introduce a model where two countries independently gather intelli-

¹For empirical evidence on transference of terrorist attacks, see Enders and Sandler (1993, 2004, 2006) and Sandler and Enders (2004). When only targets within a country is considered, terrorists’ substitution between targets is internalized by the authorities and optimal levels of counterterrorism can be reached. Powell (2007) and Bueno de Mesquita (2007) characterize optimal policies in this setting.

²Sandler and Lapan (1988, section 3) consider the effect of countries jointly becoming better informed about the terrorists, but do not study differences in information between countries, which is the focus here.
gence about the capabilities of a transnational terrorist organization, which means that each country is uncertain about the intelligence collected by the other country. Following this intelligence gathering stage, each country decides whether or not to make a costly investment in domestic (defensive) counterterrorism. Investment reduces the likelihood that an attempted terrorist attack in the homeland will be successful. Therefore, if one country invests while the other does not, the terrorists will attack the non-investing country. Otherwise the terrorists are equally likely to attack each country.

To understand the implications of national intelligence gathering, the equilibrium outcome of the model is compared to the outcome of a common intelligence benchmark model where all intelligence is known to both countries when they make their policy decisions. With common intelligence, the standard result that negative externalities lead to overprovision of domestic counterterrorism holds. Thus, a fundamental question is if national intelligence gathering leads to further inefficiencies or if it perhaps mitigates the overprovision from the benchmark model.

A main result of the paper is that if the countries’ ex ante expectation about the capabilities of the terrorists (i.e., their common expectation before they each gather new intelligence) is relatively high and if counterterrorism investments substantially decrease the success rate of attempted attacks, then national intelligence gathering increases the inefficiencies in counterterrorism provision relative to the common intelligence benchmark. Thus, in times of a substantial general threat from transnational terrorism such as the threat from Al-Qaeda and related groups in recent years, the relevant conclusion from the model is that, because intelligence is gathered on the national level, provision of (relatively effective) domestic counterterrorism will be even less efficient than what is to be expected solely because of the negative transnational externalities.

The reasons behind this finding are twofold. First, when intelligence is gathered nationally it is possible that the two countries end up with different beliefs about the terrorists’ capabilities. Thus we may get the outcome that one country invests in domestic counterterrorism while the other does not, which is not possible in equilibrium when the countries have the same information. This outcome is, efficiency wise, bad since investment by only one country simply makes the terrorists attack the non-investing country and therefore does not reduce the total damage from terrorism. Second, even when the countries make the same investment decision in the national intelligence game, the outcome is still less efficient than the benchmark. In particular, the countries will sometimes both invest when the benchmark outcome is efficient mutual non-investment. This happens because each country is uncertain about the intelligence gathered by the

3The type of counterterrorism considered can be referred to as domestic or defensive. The important characteristic is that it involves actions in the homeland that make terrorist attacks less likely to succeed there without substantially reducing the general capabilities of the transnational terrorist organization. I refer to it as domestic because it can include measures that would typically be called proactive, for example programs that seek to infiltrate terrorist cells in the homeland. Such measures make it harder for cells to fulfill their missions and thus make the terrorist organization more likely to fund or send cells elsewhere.
other country. Country $i$ will take into account that country $j$ may have received intelligence that will make it invest, which transfers terrorist activity to country $i$. This increases country $i$’s incentive to invest compared to the common intelligence benchmark.

Given this paper’s focus on national intelligence gathering, it is natural to consider the possibility of intelligence sharing between countries. There are clear benefits from mutual intelligence sharing. First of all, it helps countries make more accurate assessments of transnational terrorist threats because they receive information from a wider range of sources. Further, it allows each country’s intelligence authorities to specialize within its particular areas of strength and reduces countries’ uncertainty about the information and assessments of other countries, which makes efficient policy coordination easier. However, sharing of intelligence also involves risks. For example, it can lead to unwanted revelation of sources and capabilities, which may help adversaries and make future intelligence gathering harder. Also, increased specialization leads to higher dependency on other countries’ intelligence services and thus increased vulnerability to defection.

The relationship between the US and Pakistan serves as a good example of the dilemmas of intelligence sharing (e.g., Walsh 2009, chapter 5). For the US, the potential benefits of intelligence sharing with Pakistan are very high because Pakistani agencies are in a unique position to collect intelligence on Al-Qaeda, especially human intelligence. Further, Al-Qaeda is also a threat to the security of Pakistan, so there are also clear incentives for Pakistani authorities to cooperate with the US. Despite these potential benefits to both sides, US-Pakistani cooperation in this area has generally been difficult. Cooperation by Pakistani authorities has been limited by lack of domestic political support for cooperating with the US and some sympathy for Al-Qaeda and the Taliban within the intelligence services. The presence of these sympathies within Pakistani authorities also makes intelligence sharing dangerous for the US because of the substantial risk of leaks. This is very well documented by US behavior before the raid to kill Osama bin Laden in Pakistan on May 2, 2011. Pakistani authorities were not informed about the mission at all because of fear that this information would be leaked, which could have allowed bin Laden to escape.⁴

I explore the possibilities for intelligence sharing within the framework of this paper by extending the main model with a communication stage taking place before the investment decisions. It turns out that results on intelligence sharing are highly dependent on whether it is possible to verifiably share intelligence or not. If intelligence can be shared in a way that makes the content verifiable by the receiver, full intelligence sharing is possible in equilibrium and thus we can get the common intelligence benchmark outcome even though intelligence is gathered nationally. However, the nature of

⁴Leon Panetta, Director of the CIA during the planning and the mission, has clearly stated this in later interviews. See for example this interview with CBS News’ 60 Minutes: http://www.cbsnews.com/8301-18560_162-57367997/the-defense-secretary-leon-panetta/
intelligence and/or the relations between targeted countries and their intelligence agencies often makes verifiable sharing either completely impossible or at least a non-viable option (Walsh 2006, 2009). Therefore, I also consider a case where only cheap talk communication is possible. With this assumption, intelligence will generally not be credibly shared and thus the national gathering of intelligence is still important for the outcome.

The reasons behind the results on intelligence sharing are relatively simple. In both cases, the key observation is that each country always prefers non-investment by the other country because the terrorists are then more likely to attack there. This implies that each country prefers the other country to believe that the capabilities of the terrorists are low. With verifiable sharing, this means that a country with intelligence suggesting the terrorists are weak will reveal this information. Further, this will make the country also reveal when it expects the terrorists to be slightly stronger (to show that it does not believe them to be much stronger) and so on. Thus all intelligence will be revealed. When only cheap talk communication is possible, each country will always pretend to have intelligence showing that the terrorists are weak. This makes communication non-credible.

I have previously discussed how this paper directly builds on the literature studying counterterrorism provision in the face of transnational terrorism. Another relevant strand of literature has focused on asymmetric information between terrorist groups and governments and the signaling aspect of terrorist attacks. Important contributions to this literature include Lapan and Sandler (1993), Overgaard (1994), and Arce and Sandler (2007, 2010). In these papers, a government is incompletely informed about important characteristics of a terrorist group, for example its capabilities or whether it is politically motivated or militant. Attacks therefore serve as a costly signaling device that makes the government update its beliefs about the terrorists. The existence of (full or partial) pooling equilibria shows that the informational asymmetry will not necessarily be overcome by signaling, which highlights the importance of better intelligence about the terrorist group. Because of the government’s incomplete information about the terrorists and the importance of intelligence, this strand of literature is clearly related to the present paper. However, my focus here is very different because the main issue is differences in information between countries (because intelligence is gathered nationally) when they respond to transnational terrorism, not the informational asymmetry between one country and a terrorist group.

In a broader international relations context, this paper can be seen as a new example of how differences in information between countries or other actors can substantially influence the outcomes of international interactions. A famous demonstration of the importance of informational differences in the international arena is the formal modeling by Fearon (1995) showing that countries’ private information about resolve or capabilities can lead to bargaining breakdown and war. While the problem studied here is not as fundamental, the findings do underscore the value of introducing more realistic assumptions about information in models of specific international interactions.
The paper is organized as follows. In section 2 I present the main model. Section 3 contains the analysis of this model and an example. In section 4 I consider intelligence sharing. Finally, in section 5 I discuss the main findings and their implications.

2 The Model

Two countries, $i = 1, 2$, are facing a common threat from a transnational terrorist organization. Each country can decide whether to make a costly investment in domestic counterterrorism or not. The cost is denoted $C$. Investment reduces the vulnerability of the country to terrorism. More precisely, the counterterrorism investment reduces the probability of an attempted attack being successful from one to $p \in (0, 1)$. The countries make their decisions independently and simultaneously. Each country’s objective is to minimize its sum of expected damages from terrorism and counterterrorism costs. Thus, a central premise of the model is that national authorities do not care about the damage incurred by other countries. So, for example, the diversion of an attack to another country will be seen as a success.

The terrorists observe the investment decisions of the countries before deciding where to attack. They can launch only one attack and their objective is to maximize expected damage. Thus, if only one country invests in domestic counterterrorism then the terrorists will attack the other country because the attack is then more likely to be successful. If neither or both countries invest, the terrorists are indifferent about where to attack. I assume that they will then attack each country with probability one half. The damage from a successful attack is denoted $D$.

2.1 Complete Information

If the two countries are completely informed about $D$, the situation can be represented by the simple $2 \times 2$ game shown below.

<table>
<thead>
<tr>
<th>Country 1</th>
<th>Country 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invest ($I$)</td>
<td>Invest ($I$)</td>
</tr>
<tr>
<td>Not Invest ($N$)</td>
<td>Not Invest ($N$)</td>
</tr>
</tbody>
</table>

5 The setup also covers implementation of domestic counterterrorism policies that are not directly costly, but lead to substantial distortionary costs for the country.

6 This type of objective function for the government of a country was first used by Sandler and Lapan (1988).
This game is easily analyzed. The pure strategy Nash equilibria are

\[
(N,N) \quad \text{if} \quad D \leq 2C, \\
(I,I) \quad \text{if} \quad D \geq \frac{2C}{2-p}.
\]

The efficient outcome (lowest total sum of expected damages and costs) is \((N,N)\) if \(D \leq \frac{2C}{1-p}\) and \((I,I)\) if \(D > \frac{2C}{1-p}\). Thus the efficient Nash equilibrium is

\[
(N,N) \quad \text{if} \quad D \leq 2C, \\
(I,I) \quad \text{if} \quad D > 2C.
\]

So we see that, assuming the countries are always able to coordinate on the efficient equilibrium, the outcome will be efficient unless \(D \in (2C, \frac{2C}{1-p})\). In this region of the parameter space the game is a prisoners’ dilemma. \(I\) strictly dominates \(N\) for each country, while the efficient outcome is that neither country invests. This is a representation of the classic result that countries facing a common transnational terrorist threat will typically overinvest in defensive counterterrorism because of negative externalities: increased investment by one country makes the terrorists more likely to attack elsewhere.

### 2.2 Uncertainty and National Intelligence

In the real world, countries are unlikely to have complete information about the capabilities of transnational terrorist organizations. Intelligence is inherently noisy. Furthermore, intelligence is generally gathered by national agencies. Thus, counterterrorism authorities are typically uncertain about both the capabilities of the terrorists and the intelligence collected by other countries. Below I will analyze the implications of introducing national gathering of intelligence in the simple model presented above. I will assume that, prior to the simultaneous counterterrorism investment decisions, each country receives a private signal (i.e., the signal is not observed by the other country) about the capabilities of the terrorists, more precisely a signal correlated with \(D\). Thus the intelligence gathering process is exogenous and the investment game with nationally gathered intelligence is a static Bayesian game.\(^8\) To clarify, the timing of the game is as follows.

1. National intelligence gathering: Each country receives a private signal about the capabilities of the terrorists;

2. Counterterrorism policy decisions: Each country independently decides whether to invest in domestic counterterrorism (and pay the cost \(C\)) or not;

\(^7\)For \(D \in \left[\frac{2C}{1-p}, 2C\right]\) there is also a mixed strategy Nash equilibrium, which is straightforward to find.

\(^8\)This type of game can also be referred to as a static game of incomplete information.
3. The terrorists launch their attack and the damage $D$ is realized if the attack is successful.

I will refer to this game as the national intelligence game.

Let $d_i \in [0, \bar{d}]$ denote the signal received by country $i$. This signal is used to form a belief about the distribution of $D$ by standard Bayesian updating. In particular, after receiving the signal $d_i$, country $i$’s expectation about $D$ is $E[D|d_i]$. I assume that $E[D|d_i]$ is a continuously differentiable and strictly increasing function of $d_i$ and that the countries have the same expectation if they receive identical signals (i.e., $E[D|d_1 = d'] = E[D|d_2 = d']$ for all $d' \in [0, \bar{d}]$). It is also assumed that the expected value of $D$ given both signals, $E[D|d_1, d_2]$, is continuously differentiable and strictly increasing with respect to both $d_1$ and $d_2$ and that it responds symmetrically to these signals (i.e., $E[D|d_1 = d', d_2 = d''] = E[D|d_1 = d'', d_2 = d']$ for all $d', d'' \in [0, \bar{d}]$).

Since the signal of each country is correlated with $D$, in general the two signals are also correlated. Let $F(\cdot|\cdot)$ denote the cumulative distribution function of one country’s signal given the signal of the other country (the situation is assumed to be completely symmetric, so the conditional distribution of $d_1$ given $d_2 = d'$ is the same as the conditional distribution of $d_2$ given $d_1 = d'$). I assume that $F(d_i|d_j)$ is continuously differentiable with respect to each signal. Furthermore, it is also assumed that an increase in $d_j$ shifts the conditional distribution of $d_i$ (weakly) to the right. More precisely, if $d' < d''$ then the distribution of $d_i$ conditional on $d_j = d''$ first order stochastically dominates the distribution of $d_i$ conditional on $d_j = d'$:

$$F(d_i|d_j = d'') \leq F(d_i|d_j = d') \quad \text{for all } d_i \in [0, \bar{d}].$$

Finally, I will also make the following minor technical assumption. Let $d'$ be a signal such that $E[D|d_i = d'] > E[D]$ ($E[D|d_i = d'] < E[D]$), where $E[D]$ denotes the ex ante expected value of $D$. Then it is assumed that

$$E[D|d', d'] > E[D|d_i = d'],$$

$$(E[D|d', d'] < E[D|d_i = d']),$$

where $E[D|d', d']$ is convenient shorthand notation for $E[D|d_1 = d', d_2 = d']$. Thus, if the reception of a particular signal makes a country update its expected value of $D$ in one direction, then learning that the other country has received the same signal will make the country update its expected value of $D$ further in the same direction.

In the national intelligence game, each country can condition its decision whether to invest or not only on its own signal. We write the strategy of country $i$ as $s_i(d_i)$. Since investment is more attractive the more damage the terrorists are capable of causing, it is natural to look for (Bayesian Nash) equilibria where each country invests if and only if its signal is above some cutoff value. Furthermore, since the game is symmetric, we
will restrict attention to symmetric cutoff equilibria.\footnote{The game is symmetric from an ex ante point of view because the countries are facing identical situations before they each receive their private signal. Of course, even though the game is symmetric, there may well exist asymmetric equilibria. However, if a symmetric equilibrium exists (which we will later show to be the case), then it is reasonable to expect the players (countries) to coordinate on such an equilibrium (e.g., Kreps 1990).} Thus, an equilibrium is given by a cutoff signal $x$ such that each country will invest if and only if its signal is above $x$:

$$s^*_i(d_i) = \begin{cases} N & \text{if } d_i \leq x \\ I & \text{if } d_i > x \end{cases}, \quad i = 1, 2. \quad (1)$$

In equilibrium it should be optimal for each country to use this strategy given that the other country is using it. Note that even though we restrict attention to symmetric equilibria, the countries may well end up making different investment decisions. For example, if country 1 receives a signal $d_1 < x$ while country 2 receives a signal $d_2 > x$ then only country 2 will invest.

## 3 Analysis of the National Intelligence Game

Before analyzing the national intelligence game, I will first consider the natural benchmark model, namely the game where the two signals (i.e., all pieces of intelligence) are commonly known by the two countries. I will refer to this game as the common intelligence (benchmark) game.

In the common intelligence game each country can condition its investment decision on both signals, so the strategy of country $i$ is written $s_i(d_1, d_2)$. This game is closely related to the complete information game because the two countries have exactly the same information available when making their (independent and simultaneous) investment decisions. Indeed, it is easy to see that the efficient equilibrium is given by the strategies

$$s^*_i(d_1, d_2) = \begin{cases} N & \text{if } E[D|d_1, d_2] \leq 2C \\ I & \text{if } E[D|d_1, d_2] > 2C \end{cases}, \quad i = 1, 2. \quad (2)$$

So the benchmark to which the outcome of the national intelligence game should be compared is that each country will invest precisely if the expected value of $D$ given all available intelligence is higher than $2C$.

To analyze the national intelligence game, assume that country $j$ is using the strategy given by the cutoff signal $x$. That is, country $j$ invests if and only if $d_j > x$. Then country $i$’s sum of expected terrorism damages and counterterrorism costs if it invests is

$$C + \frac{p}{2}(1 - F(x|d_i))E[D|d_i, d_j > x].$$

Note that $E[D|d_i, d_j > x]$ is country $i$’s expected value of $D$ conditional on its own signal $d_i$ and on country $j$ receiving a signal that is above the cutoff ($d_j > x$). If country
i chooses not to invest then its expected damage is
\[
\frac{1}{2} F(x|d_i) E[D|d_i, d_j \leq x] + (1 - F(x|d_i)) E[D|d_i, d_j > x]
\]
\[
= \frac{1}{2} E[D|d_i] + \frac{1}{2} (1 - F(x|d_i)) E[D|d_i, d_j > x].
\]
Thus it is optimal for country i not to invest precisely if
\[
\frac{1}{2} E[D|d_i] + \frac{1}{2} (1 - F(x|d_i)) E[D|d_i, d_j > x] \leq C + \frac{p}{2} (1 - F(x|d_i)) E[D|d_i, d_j > x],
\]
which is equivalent to
\[
E[D|d_i] + (1 - p)(1 - F(x|d_i)) E[D|d_i, d_j > x] \leq 2C. \tag{3}
\]
Both terms on the left hand side are increasing and continuous in \(d_i\) (by earlier assumptions, \(E[D|d_i]\) and \(E[D|d_i, d_j > x]\) are continuous and strictly increasing in \(d_i\) and \(F(x|d_i)\) is continuous and weakly decreasing in \(d_i\)). Thus it is a best response for country i to use the cutoff strategy given by \(x\) precisely if
\[
E[D|d_i] = x + (1 - p)(1 - F(x|d_i)) E[D|d_i = x, d_j > x] = 2C. \tag{4}
\]
So if \(x\) satisfies this equation then we have an equilibrium. If \(E[D|d_i = 0] < \frac{2C}{2-p}\) and \(E[D|d_i = d] > 2C\) then there exists a solution \(x > 0\) to the equation. In the following it is assumed that these assumptions hold. While we do not necessarily have uniqueness of equilibrium, it is easy to establish that a highest cutoff equilibrium exists and that this is the efficient equilibrium. All these results are proved in the appendix.

From equation (4) it immediately follows that \(E[D|d_i = x] < 2C\). Recall that in the benchmark outcome each country invests precisely if \(E[D|d_1, d_2] > 2C\). Thus, in any equilibrium of the national intelligence game each country will switch from non-investment to investment at a lower expectation about the capabilities of the terrorists than in the common intelligence benchmark game. The intuition behind this result is simple and illuminating. In the common intelligence game the countries can coordinate on not investing as long as the common expectation about \(D\) is such that \(N\) is the best response to \(N\) in the complete information game. In the national intelligence game each country is uncertain about the intelligence gathered by the other country. Thus there is always a positive probability that the other country’s expected value of \(D\) will make it invest. This means that, even when country i’s expectation about \(D\) is such that the countries could coordinate on mutual non-investment if that expectation had been common for the two countries, it can be optimal for country i to invest. Below I formulate the observation as a proposition.

**Proposition 1** In any equilibrium of the national intelligence game, each country will switch from non-investment to investment in domestic counterterrorism at lower expected capabilities of the terrorists than in the efficient equilibrium of the common intelligence benchmark game.
While this result is interesting, it does not imply that national intelligence gathering necessarily worsens the inefficiencies in provision of domestic counterterrorism relative to the benchmark. Suppose both countries have received the cutoff signal. Then each country’s expected value of $D$ is $E[D|x] < 2C$ (for simplicity we write $E[D|x]$ instead of $E[D|d_i = x]$). However, it could be that the expected value of $D$ given the information that both signals is equal to $x$, $E[D|x, x] = E[D|d_1 = x, d_2 = x]$, is higher than $2C$. If this is the case then there exist signal pairs such that national intelligence gathering leads to efficient mutual non-investment while the countries would have both invested if all intelligence had been commonly known. In the following I will thoroughly analyze when national intelligence gathering does in fact lead to more inefficient provision of domestic counterterrorism.

As the comments above suggest, the comparison of the outcomes of the national intelligence game and the common intelligence benchmark game depends critically on whether $E[D|x, x] < 2C$ or $E[D|x, x] > 2C$ (for simplicity we ignore the boundary case). Examples of these two cases are shown in figure 1 and 2. Remember that in the efficient equilibrium of the benchmark game the countries both invest precisely if the expected value of $D$ given both signals is above $2C$. The efficient outcome is for both countries to invest if $E[D|d_1, d_2] > \frac{2C}{1-p}$ and for none of them to invest otherwise.

Figure 1: $E[D|x, x] < 2C'$
If $E[D|x, x] < 2C$ then the set of signal pairs such that neither country invests in the national intelligence game is contained in the corresponding set for the benchmark game. So if we only consider signal pairs where the countries make the same investment decision in the national intelligence game then the national intelligence outcome is, efficiency wise, worse than the common intelligence benchmark outcome. However, for the benchmark outcome to be at least as good as the national intelligence outcome for all signal pairs we need the restriction $p \leq \frac{1}{2}$ (see the appendix for a proof of this claim). The problem occurs when only one country invests in the national intelligence game and $E[D|d_1, d_2] \in (2C, \frac{2C}{1-p})$. For such signal pairs the common intelligence outcome, $(I, I)$, is only better than one country investing if $p \leq \frac{1}{2}$, i.e., if investment in domestic counterterrorism reduces the probability that an attempted attack is successful by at least fifty percent.

Now suppose $E[D|x, x] > 2C$. Then, if we are in the national intelligence game and both countries have received the cutoff signal, each country’s expected value of $D$ is below $2C$, but the expected value given all information in the game is above $2C$. This implies that there exist signal pairs such that national intelligence gathering leads to efficient mutual non-investment by the two countries, while the outcome with commonly known intelligence would be mutual investment. More precisely, this is the case for signal pairs with $d_1, d_2 \leq x$ and $E[D|d_1, d_2] \in (2C, \frac{2C}{1-p})$.

If $E[D|x, x] > \frac{2C}{1-p}$ then national intelligence leads to inefficient mutual non-investment for signal pairs with $d_1, d_2 \leq x$ and $E[D|d_1, d_2] > \frac{2C}{1-p}$ since $(I, I)$ is the efficient outcome when the expected value of $D$ is above $\frac{2C}{1-p}$.

Finally, when we consider signal pairs with $d_i < x < d_j$ such that one country invests while the other does not, then there always (no matter if $E[D|x, x]$ is above or below
exist signal pairs such that the benchmark is better than the national intelligence outcome. For example, suppose \( d_1 < x, d_2 > x \), and \( E[D|d_1, d_2] < 2C \).\(^{10}\) Then the outcome of the national intelligence game is \((N, I)\) while the benchmark outcome is \((N, N)\). And \((N, N)\) always leads to a lower total sum of damages and costs than investment by one country since the investment then is, from an efficiency point of view, wasted because the terrorists will attack the other country.

Thus, to sum up, when \( E[D|x, x] < 2C \) and \( p \leq \frac{1}{2} \) we have the unambiguous result that national intelligence gathering makes the provision of domestic counterterrorism more inefficient. When \( E[D|x, x] > 2C \) or \( p > \frac{1}{2} \) the efficiency implications of national intelligence gathering are ambiguous: for some signal pairs it leads to worse outcomes, for others it is the other way around. The precise results are collected in the following proposition.

**Proposition 2** Let \( x \) be an equilibrium of the national intelligence game. Then the following statements hold.

1. If \( E[D|x, x] < 2C \) then, for all signal pairs such that the national intelligence outcome is \((N, N)\) or \((I, I)\), the benchmark outcome results in a weakly lower total sum of counterterrorism costs and terrorism damages (strictly lower for some signal pairs). If \( p \leq \frac{1}{2} \) this holds for all signal pairs.

2. If \( E[D|x, x] \in (2C, \frac{2C}{1-p}) \) then, for all signal pairs such that the national intelligence outcome is \((N, N)\) or \((I, I)\), the benchmark outcome results in a weakly higher total sum of counterterrorism costs and terrorism damages (strictly higher for some signal pairs). However, there exist signal pairs such that the national intelligence outcome is \((N, I)\) or \((I, N)\) and the benchmark outcome results in a strictly lower total sum of costs and damages.

3. If \( E[D|x, x] > \frac{2C}{1-p} \) then there exist signal pairs such that the national intelligence outcome is \((N, N)\) and the benchmark outcome results in a strictly higher total sum of counterterrorism costs and terrorism damages. However, there also exist signal pairs with the same national intelligence outcome (and with the outcomes \((N, I)\) and \((I, N)\)) such that the benchmark outcome results in a strictly lower total sum of costs and damages.

While the proposition reveals how the comparison between the national intelligence outcome and the common intelligence benchmark depends on the expected capabilities of the terrorists conditional on both countries receiving the cutoff signal, it does not directly

\(^{10}\)Such signal pairs exist. First note that, because 0 is the lowest possible signal, we have \( E[D|d_1 = 0, d_2 = x] < E[D|x] < 2C \). Therefore, by choosing \( d_1 < x \) sufficiently close to zero and \( d_2 > x \) sufficiently close to \( x \) we get a signal pair with \( E[D|d_1, d_2] < 2C \) (by assumption, \( E[D|d_1, d_2] \) is continuous with respect to each of the signals).
reveal how the comparison depends on the primitives of the model. Most importantly, under which conditions does it hold that \( E[D|x, x] < 2C? \)

Since \( E[D|x] < 2C \), it immediately follows that \( E[D|x, x] < 2C \) if \( E[D|x, x] \leq E[D|x] \). By an assumption made earlier, this is the case if \( E[D|x] < E[D] \). So if \( E[D] \geq 2C \) then we are in case one of Proposition 2. Thus we have the following important result.

**Proposition 3** Suppose the common ex ante expected value of \( D \) is above the level where it makes investment a strictly dominating strategy for each country if no further information is received (\( E[D] \geq 2C \)). Then national intelligence gathering makes the provision of domestic counterterrorism more inefficient relative to the common intelligence benchmark.

If \( E[D] < 2C \) then it is possible that we are in case two or three of Proposition 2. Below I present an example of the general model studied so far. The example shows that it is indeed possible for national intelligence gathering to lead to more efficient outcomes than if all intelligence is commonly known by the two countries. However, for this to happen \( E[D] \) has to be substantially below \( 2C \), especially when \( p \) is relatively small (i.e., when investment in counterterrorism is quite effective in reducing a country’s vulnerability to a terrorist attack).

### 3.1 An Example

Suppose the two countries independently collect intelligence about different parts of the transnational terrorist organization’s activities. For example, it could be that each country primarily collects intelligence in a specific geographical region because their intelligence network is most developed there for historical or other reasons. Further, assume that each country learns everything about the particular activities it surveils and that no activities escape scrutiny, such that the true capabilities of the entire terrorist organization is simply the sum of the signals received by the two countries:

\[
D = d_1 + d_2.
\]

Thus, I basically assume away the inherent noisiness of intelligence. A country is only uncertain about the capabilities of the terrorists because it does not have information about the activities surveilled by the other country. The signals \( d_1 \) and \( d_2 \) are assumed to be independent and uniformly distributed on \([0, 1]\). While this is clearly a stylized example, it serves the purpose of illustrating the mechanisms at play in the model, for example how national intelligence gathering can, under some circumstances, make the provision of domestic counterterrorism more efficient.

In this example, the benchmark outcome is \((N, N)\) if \( d_1 + d_2 \leq 2C \) and \((I, I)\) otherwise. Note that the benchmark game is identical to the complete information game because the true value of \( D \) is known if both signals are known.
In the game with national intelligence gathering, the equilibrium equation (4) becomes
\[ x + \frac{1}{2} + (1 - p)(1 - x)(x + \frac{1 + x}{2}) = 2C. \]
This is simply a second order equation in \( x \):
\[ 3(1 - p)x^2 - 2(2 - p)x - (2 - p) + 4C = 0. \] (5)
The assumptions made earlier that \( E[D_i = 0] < \frac{2C}{2 - p} \) and \( E[D_i = \bar{d}] > 2C \) become
\[ \frac{1}{2} < \frac{2C}{2 - p} \quad \text{and} \quad 1 + \frac{1}{2} > 2C, \]
which is equivalent to
\[ \frac{2 - p}{4} < C < \frac{3}{4}. \]
Here these assumptions imply that there is a unique equilibrium \( x \in (0, 1) \).\(^{11}\) This equilibrium is found by solving equation (5) for the smallest root (with the assumptions made, the polynomial on the left hand side is positive at \( x = 0 \) and negative at \( x = 1 \)). The exact expression for \( x \) can be found in the appendix.

Having found the equilibrium \( x \) in the national intelligence game, we are now ready to analyze when national intelligence gathering generally makes provision of domestic counterterrorism less efficient than in the benchmark game and when the opposite is true for some signal pairs. From the general theory we know that if \( E[D] \geq 2C \) then national intelligence gathering always leads to (weakly) worse outcomes. Here \( E[D] = E[d_1] + E[d_2] = 1 \), so the condition \( E[D] \geq 2C \) corresponds to \( C \leq \frac{1}{2} \). Therefore, we are primarily interested in analyzing what happens when \( \frac{1}{2} < C < \frac{3}{4} \).

First consider the special case \( p = \frac{1}{2} \). From the formula in the appendix we get that the equilibrium is
\[ x = 1 - \sqrt{2 - \frac{8}{3}C}. \]
So \( E[D|x, x] = x + x = 2 - 2\sqrt{2 - \frac{3}{8}C} \). This means that we are in case one of Proposition 2 precisely if
\[ 2 - 2\sqrt{2 - \frac{8}{3}C} < 2C, \]
i.e., if
\[ C < -\frac{1}{3} + \frac{1}{6}\sqrt{40} \approx 0.72. \]
So, efficiency wise, national intelligence gathering is unambiguously worse than the benchmark even for values of \( C \) such that the ex ante expected value of \( D \) is substantially
\(^{11}\)If \( C \leq \frac{2-p}{2} \) then we still have a unique equilibrium: \( x = 0 \). If \( C \geq \frac{3}{4} \) then \( x = 1 \) is an equilibrium. And while there can be other equilibria, \( x = 1 \) is the efficient one.
below $2C$. But if, approximately, $C \in (.72, .75)$ then we are in case two of Proposition 2 (it is easy to check that we are not in case three).

In figure 3, $E[D|x, x] - 2C$ is plotted as a function of $C$ for different values of $p$ ($p = .1, .25, .5, .75, \text{and} .9$). Thus, for each value of $p$ it is easy to see when $E[D|x, x]$ is below $2C$, such that we can use Proposition 2 to conclude that national intelligence gathering makes provision of domestic counterterrorism less efficient than in the benchmark model. We see that the smaller $p$ is, the larger is the region where $E[D|x, x] < 2C$. This observation is formally proved in the appendix. Furthermore, for $p$ sufficiently small, the benchmark outcome is better than the national intelligence outcome for all $C$’s.

Figure 3: $E[D|x, x] - 2C$ as a function of $C$ for different values of $p$

To sum up, this example has illustrated that it is possible for national intelligence gathering to mitigate the standard inefficiencies in provision of domestic counterterrorism (overprovision due to negative externalities). If the ex ante expectation about the capabilities of the terrorists is very low relative to $2C$, then it is only if the observed intelligence greatly increases the expected value of $D$ that a country will choose to invest. Such a strong adjustment of expectations can happen for lower signals if we are in the benchmark game where two signals are received than if each country only observes its own signal. And this effect can dominate the effect deriving from the fact that the adjustment of expectation needed to make a country invest is lower in the national intelligence game (because of uncertainty about the intelligence collected by the other country, see Proposition 1 and the discussion before it). When this is the case, there exist signal pairs such that the national intelligence outcome is efficient mutual non-investment while the benchmark outcome is inefficient mutual investment.

We also saw that the smaller $p$ is, the larger is the set of investment costs for which the benchmark is generally more efficient than the national intelligence outcome. The intuition behind this result is worth discussing. For all $p < 1$, the reduction in expected domestic damage for a country if it invests is higher when the other country invests
than when it does not. However, the difference in reduction is larger the smaller $p$ is. This means that the smaller $p$ is the more it matters that, with national intelligence gathering, there is always a positive probability that the other country will invest. And this implies that national intelligence gathering generally increases the inefficiencies in counterterrorism provision for a larger set of $C$’s when $p$ is smaller.

Finally, it is important to emphasize that if investment in counterterrorism is reasonably effective in reducing a country’s vulnerability to a terrorist attack, national intelligence gathering generally leads to less efficient outcomes unless the ex ante expected value of $D$ is very low relative to $2C$. So, in times of a substantial general terrorist threat such as the threat from Al-Qaeda and related groups in recent years, the relevant implication of the model is that, because intelligence is gathered on the national level, provision of domestic counterterrorism will be even less efficient than what is to be expected solely because of the negative transnational externalities.

4 Intelligence sharing

Until now I have assumed that national gathering of intelligence implies that each country only has access to its own intelligence (signal) when deciding whether to invest in domestic counterterrorism or not. However, it could be that the countries manage to credibly share their intelligence. If so, the observation that intelligence is gathered by national agencies is clearly less relevant for counterterrorism provision.

In this section I explore the problem of intelligence sharing. I do so by extending the national intelligence game studied above with a communication stage taking place before the investment decisions. In this stage, each country has the option of sending a message to the other country or not. The possible messages are the possible signals that the countries can receive. If each country always sends a message equal to its received signal then all intelligence become commonly known (by Bayesian updating) and the subsequent investment game will then be equivalent to the benchmark game. If each country never sends a message (or, for example, always sends the message $m = 0$), no information is revealed.

I will distinguish between two cases with respect to the possibilities for intelligence sharing. In the first case it is assumed that the messages are pure cheap talk. Thus a country can send any message at no cost and the receiving country has no way of verifying if a message is correct. The nature of intelligence implies that it is often not possible to directly communicate the content of a particular piece of intelligence in a verifiable way. For example, information that one country has obtained by interviewing informants can typically not be shared in a way that makes the receiving country certain that the content has not been manipulated. Furthermore, intelligence agencies generally have strong incentives not to reveal their sources and other details about their activities. Thus, even when some piece of intelligence can in principle be shared verifiably, doing so may well not be seen as a viable option because this would reveal, for example, the
sources from which the intelligence was obtained or specific capabilities of the agency.

The second case corresponds to a situation where the content of a message is verifiable. I assume that each country $i$ can either send the true message $m_i = d_i$ (and country $j$ will then know the signal of country $i$ with certainty) or it can choose not to send a message.$^{12}$ This case could be relevant when there is a large degree of trust between the intelligence agencies of the two countries.$^{13}$ With a high level of trust, it is likely that authorities are less reluctant to reveal the highly classified information necessary for verifiable intelligence sharing because the risk that it will be passed on to third parties is perceived to be low. Furthermore, even when the intelligence collected is of a kind that makes direct verifiable sharing difficult, large potential reputation costs from sending a false message to a trusted partner implies that this could still be the relevant case, at least as an approximation. Thus, to give an example, the cheap talk case is likely to be more relevant if the two countries are the US and Pakistan while the verifiable messages case is more relevant if they are the US and the UK, which is a good example of two countries whose intelligence authorities have built up mutual trust through a close long term relationship, especially during the cold war (e.g., Walsh 2009, chapter 2).

Our analysis of the two cases is presented below. The models with communication are quite complex (especially because communication is two-sided and the set of possible signals is continuous), so the formal analysis will focus on equilibria with full or no revelation of intelligence between the countries. In the cheap talk game, full revelation is not possible in equilibrium while equilibria with no revelation do exist. For the game with verifiable messages it is the other way around. A complete equilibrium analysis of the two games is beyond the scope of this paper, but the arguments behind the results on the two extreme types of equilibria (full or no revelation) provide powerful intuition about the barriers to credible intelligence sharing when only cheap talk communication is possible and how the mere possibility of sending verifiable messages facilitates actual sharing.

### 4.1 Cheap Talk Communication

Suppose first that there exists a (Perfect Bayesian) equilibrium with full revelation of each country’s signal. With full revelation, the investment game is equivalent to the benchmark game studied earlier and it is assumed that the countries will then play the efficient benchmark equilibrium, i.e., the outcome will be $(N, N)$ if $E[D|d_1, d_2] \leq 2C$ and $(I, I)$ otherwise. For simplicity, assume that each country always sends the true message $m_i = d_i$. The possibility of sending unverifiable cheap talk messages could have been included in this case, but this would not change the results.

$^{12}$The possibility of sending unverifiable cheap talk messages could have been included in this case, but this would not change the results.

$^{13}$Even when two countries are close allies, it does not necessarily imply that there is a high level of trust between their intelligence agencies. For example, the "Curveball" case related to the 2003 invasion of Iraq illustrates that there was substantial distrust between the intelligence agencies of Germany and the US (Drogin 2007).
message, i.e., the equilibrium message of country $i$ is $m^*_i(d_i) = d_i$, $i = 1, 2$.\footnote{This is of course only one possible equilibrium with full revelation, but the argument below holds generally.} Consider the decision of country 1 in the message stage when it has received a signal $d_1 = d' > 0$. In equilibrium, country 1 sends the true message $m_1 = d'$. However, suppose the country deviates and instead sends the message $m_1 = 0$. Then, in the investment stage, country 2 will believe that country 1 has received the signal $d_1 = 0$ and therefore country 2’s expectation about $D$ will be $E[D|d_1 = 0, d_2]$, whereas it would have been higher ($E[D|d_1 = d', d_2]$) if country 1 had not deviated. Thus, the deviation by country 1 makes country 2 invest for a strictly smaller range of realized $d_2$’s. And since it is always better for country 1 that country 2 does not invest (because the terrorists are then more likely to attack there), this means that it is better for country 1 to send the false message $m_1 = 0$ rather than the true message $m_1 = d'$, which contradicts the equilibrium assumption. In other words, full intelligence revelation is not sustainable because it implies that each country always has an incentive to make the other country less likely to invest by pretending to have received a lower signal about the terrorists’ capabilities than it really has.

Essentially the same argument as above suggests that there can be no useful revelation at all in equilibrium. More specifically, if we have an equilibrium where some country $i$ sends different messages ($m'_i$ and $m''_i$) for two different signals ($d'_i$ and $d''_i$), then the investment behavior of country $j$ does not depend on which of these messages it receives. Suppose it did. Then country $i$ would, after receiving any of the two signals, have an incentive to send the message that minimizes the probability of investment by country $j$. Assume that this is the message $m'_i$. Then, after receiving the signal $d''_i$ country $i$ would have an incentive to deviate to sending the message $m''_i$. It should be emphasized that this argument is not a formal proof. A formal proof would require considering all possible equilibria with any useful revelation, which would be a huge task in a complex game like this (for example it would be necessary to consider investment situations with all kinds of asymmetric beliefs for the two countries). However, the argument clearly illustrates the barriers to any credible intelligence sharing through cheap talk communication.

Finally, it is easy to show that there exist equilibria with no revelation of information, such that the investment stage is equivalent to the national intelligence game. Suppose for example that each country never sends a message and then, in the investment stage, plays the equilibrium strategy with the highest cutoff signal from the national intelligence game. Then we just have to check that it is never profitable for a country to actually send some message in stage one. This is the case if the out of equilibrium belief of each country after seeing any message is that the sender country has received the highest possible signal $\bar{d}$ (which will make the receiving country maximally likely to invest). Because then it is obviously better to send no message, which reveals no information to the other country. It is also possible to simply let the out of equilibrium
belief of country \( j \) after receiving a message from country \( i \) be the same as the belief after receiving no message, i.e., the belief from the national intelligence game (which is of course given by the conditional distribution of \( d_i \) given \( d_j \), \( F(d_i|d_j) \)). Then, if a country deviates to sending some message instead of no message it does not change the outcome, so we still have an equilibrium.

### 4.2 Verifiable Sharing

I first show that there exists an equilibrium with full revelation. Let \( m_i^*(d_i) = d_i \), \( i = 1, 2 \). Further, let the countries’ behavior in the investment stage (where both signals are commonly known because of the messages sent) be as in the efficient equilibrium of the common intelligence benchmark game. I.e., the countries’ investment strategies are given by equation (2) in section 3. Then we just have to check that each country will never find it optimal to deviate by not sending a message. This of course depends on the belief of the other country if it receives no message, which is an out of equilibrium belief. Let this belief be that the signal of a country sending no message is the highest possible signal \( \tilde{d} \).\(^{15}\) With this out of equilibrium belief, neither country can ever profitably deviate from revealing its signal, because such a deviation would (weakly) increase the probability that the other country invests. So when intelligence can be shared verifiably, it is possible to reach the common intelligence benchmark outcome even though intelligence is gathered on the national level.

To see why no intelligence revelation is not possible in equilibrium, assume that we have such an equilibrium. Then the outcome will be identical to the outcome of the national intelligence game. But this means that if country \( i \) has received a very low signal, for example \( d_i = 0 \), then it has an incentive to deviate from sending no message to verifiable sharing of its signal. Because this will make country \( j \) invest for a strictly smaller range of \( d_j \)'s, which makes country \( i \) better off.

A somewhat informal extension of the argument above suggests that all useful intelligence will be shared when verifiable messages are possible. Each country always has an incentive to reveal very low signals, which further means that there is an incentive to reveal slightly higher signals, and so on. This unraveling may stop if we reach a point where, for any higher signal, the country will always invest in counter-terrorism. This is the case if there exist signals \( d_i \) such that \( E[D|d_i, d_j = 0] > 2C \). Because then country \( i \) does not have an incentive to reveal that it has received a signal \( d' \) with \( E[D|d_i = d', d_j = 0] > 2C \) rather than an even higher signal since investment will be a dominating strategy in either case, which means that revelation will not influence the decision of country \( j \). Still, because of the unraveling, all signals \( d_i \) with

\(^{15}\)With this belief, a country that has received the signal \( \tilde{d} \) is obviously indifferent between revealing its signal and sending no message. Thus, sending no message is not equilibrium dominated for a "type \( \tilde{d} \) country" and therefore the equilibrium is not in conflict with standard refinement procedures à la the Intuitive Criterion (Cho and Kreps 1987).
$E[D|d_i, d_j = 0] \leq 2C$ will be directly revealed\textsuperscript{16}, which means that the countries can reach the benchmark outcome. While this is not a complete formal proof, the argument clearly illustrates how the possibility of direct verifiable sharing of intelligence creates the incentives that lead to actual sharing.

\section{Discussion}

In this paper I have explored the implications of national intelligence gathering for the provision of domestic (defensive) counterterrorism when two countries are facing a common transnational terrorist threat. An important finding was that, unless the ex ante expectation about the capabilities of the terrorists is quite low (or counterterrorism is relatively ineffective), national gathering of intelligence generally increases the inefficiencies in counterterrorism provision relative to a benchmark model where all collected intelligence is known to both countries. Thus, this is clearly the relevant implication of the model in times of a substantial general threat from transnational terrorist organizations, for example the period since the existence, intentions, and general capabilities of Al-Qaeda and related groups became known in the 1990s. Policy wise, the model thus highlights particular benefits of increased international cooperation on intelligence matters. More cooperation has the potential to reduce both the uncertainty about the threats from transnational terrorism and the uncertainty about the intelligence collected by other countries. This can, in turn, lead to better policy coordination, although the negative externalities due to transference of terrorist attacks will of course still be a problem. The externality problems are hard to overcome because countries are generally very reluctant to give up autonomy over security policy, but greater economic and political integration (for example in the European Union) may over time lead to some internalization of the costs that domestic counterterrorism measures impose on other countries.

In the last part of the paper I studied the problem of intelligence sharing within the modeling framework used earlier. It was shown that the level of actual credible sharing depends critically on whether it is possible for each country to verifiably share its intelligence or not. If verifiable sharing is truly possible, intelligence will indeed be shared and the counterterrorism investment outcome will be similar to that of the benchmark model with commonly known intelligence. However, if the nature of the intelligence gathered or the lack of trust between intelligence agencies implies that only cheap talk communication is possible between the countries, then intelligence will not be credibly shared and thus the national intelligence gathering still matters for the outcome. While the two cases considered (fully verifiable messages and pure cheap talk) are clearly stylized, the findings do highlight the importance of trust for true, useful intelligence sharing. If there is enough trust that verifiable sharing is really viewed as

\textsuperscript{16}Thus, if country $j$ does not receive a signal it can infer that $E[D|d_i, d_j = 0] > 2C$. 

an option when practically possible (because the risk of misuse, leaks to third parties, etc. is perceived to be low), then this can in itself lead to actual sharing, even when countries have somewhat divergent policy interests because they each try to minimize the damage from terrorism (and counterterrorism costs) in the homeland. Since trust is built gradually through long term relationships, this also implies that investment in developing such relationships may well be worthwhile even if the short term benefits seem limited.
References


Appendix

Claims made in section 3, immediately after equation (4).

1. There exists an equilibrium $x > 0$.
   
   **Proof.** Let $G : [0, \bar{d}) \rightarrow \mathbb{R}$ be the function defined by
   
   $$G(d) = E[D|d_i = d] + (1-p)(1-F(d|d))E[D|d_i = d, d_j > d] - 2C$$
   
   for all $d \in [0, \bar{d})$.

   From our assumptions it follows that $G$ is continuous and that
   
   $$G(0) = E[D|d_i = 0] + (1-p)E[D|d_i = 0] - 2C < (2-p)\frac{2C}{2-p} - 2C = 0$$

   and
   
   $$G(d) > E[D|d_i = d] - 2C > 0$$
   
   for $d$ sufficiently close to $\bar{d}$.

   From these observations it follows that there must exist an $x \in (0, \bar{d})$ with $G(x) = 0$, which means that equation (4) is satisfied. □

2. There exists a highest cutoff equilibrium and this is the efficient equilibrium.
   
   **Proof.** By continuity of $G$ (see claim 1 above), the set of equilibrium cutoff signals, $\{x | G(x) = 0\}$, is closed and bounded in $\mathbb{R}$. Thus there exists a highest cutoff equilibrium $x^H \in (0, \bar{d})$.

   To see that $x^H$ is the efficient symmetric cutoff equilibrium, let $u_i(x_i, x_j)$ denote the ex ante expected utility of country $i$ when it uses the cutoff strategy given by $x_i$ and country $j$ uses the cutoff strategy given by $x_j$. It is easy to see that country $i$ is always better off when country $j$ is less likely to invest, i.e., when $x_j$ is higher. Thus, if $x \neq x^H$ is a symmetric cutoff equilibrium, then we have
   
   $$u_i(x, x) < u_i(x, x^H).$$

   Furthermore, since $x^H$ is a symmetric equilibrium cutoff, we also have
   
   $$u_i(x, x^H) \leq u_i(x^H, x^H).$$

   Thus, each country $i$ receives a higher ex ante expected utility in the $x^H$ equilibrium than in any other symmetric cutoff equilibrium. □

Claim made in section 3, between Proposition 1 and 2.

Suppose $E[D|x, x] < 2C$. Then, efficiency wise, the benchmark outcome is (weakly) better than the national intelligence outcome for all realized signal pairs precisely if $p \leq \frac{1}{2}$. 

25
Proof. In the main text it is argued that the benchmark outcome is (weakly) better than
the national intelligence outcome if \(d_1, d_2 \leq x\) or \(d_1, d_2 > x\). Further, if \(E[D|d_1, d_2] \leq 2C\) or \(E[D|d_1, d_2] \geq \frac{2C}{1-p}\), then the benchmark outcome is efficient and therefore obviously at
least as good as the national intelligence outcome. Thus the only remaining cases are
signal pairs such that precisely one country invests in the national intelligence game and
\(E[D|d_1, d_2] \in (2C, \frac{2C}{1-p})\) (which implies that the benchmark outcome is \((I, I)\). \((I, I)\) is
at least as good as \((N, I)\) (and \((N, I)\)) precisely if
\[
2C + pE[D|d_1, d_2] \leq C + E[D|d_1, d_2]
\]
(the left hand side is the total sum of damages and costs if both countries invest, the
right hand side is the same sum when only one country invests). This is equivalent to
\[
E[D|d_1, d_2] \geq \frac{C}{1-p}.
\]
Thus mutual investment is better than only one country investing for all signal pairs
with \(E[D|d_1, d_2] \in (2C, \frac{2C}{1-p})\) precisely if \(\frac{C}{1-p} \leq 2C\). This is equivalent to \(p \leq \frac{1}{2}\). □

The unique equilibrium in the example in section 3.1.
Solve the second order equation (5) to get that the smallest root (and thus the equilib-
rium cutoff signal) is
\[
x = \frac{(2 - p) - \sqrt{(2-p)^2 - 3(1-p)(-2 + p + 4C)}}{3(1-p)}.
\]

Claim made in section 3.1.
The smaller \(p\) is, the larger is the set of \(C\)’s such that \(E[D|x, x] < 2C\).
Proof. First note that \(E[D|x, x] = x + x = 2x\). Thus \(E[D|x, x] < 2C\) is equivalent to
\(x < C\). Therefore it suffices to show that \(\frac{\partial x}{\partial p} > 0\) for all \(p \in (0, 1)\) and \(C \in \left(\frac{2-p}{4}, \frac{3}{4}\right)\).
From the expression for \(x\) from above we get
\[
\frac{\partial x}{\partial p} = 3 \frac{\frac{1}{2}(1-p)(13 - 8p - 12C)A^{\frac{1}{2}} - A^{\frac{1}{2}}}{9(1-p)^2},
\]
where \(A = (2-p)^2 - 3(1-p)(-2 + p + 4C)\). So it follows that \(\frac{\partial x}{\partial p} > 0\) is equivalent to
\[
A^{\frac{1}{2}} + \frac{1}{2}(1-p)(13 - 8p - 12C) - A > 0.
\]
By further calculations this is equivalent to
\[
144(1-p)^2C(\frac{5}{6} - C) > 9(1-p)^2,
\]
which is easily seen to hold for all \(p \in (0, 1)\) and \(C \in \left(\frac{2-p}{4}, \frac{3}{4}\right)\). □