For Online Publication: Self-Managing Terror Appendix

Peter Schram

January 11, 2021

Abstract

This Online Appendix is divided into four parts. Part I discusses the set of cases of radical Islamist insurgencies where this theory could apply. Part II describes the full equilibrium strategies across the four techniques that the principal uses. Part II also more fully describes the Heterogeneous Teams with Incentive Contracts technique. Part III provides proofs for Propositions 1-4 and Lemma 1. Part IV provides a more detailed discussion on Observations 1 and 4. Part V describes what occurs when the principal introduces the "Perfectly Aligned" agent and the other extensions. As a note, I consider a more-parameterized version of the agents' utility functions relative to the functional form that I selected in the paper; I discuss this change at the onset of Part II.

Contents

I Substantive Importance 1

1 Groups utilizing radical Islamist Foreign Fighters 1

II Full Equilibrium Strategies 4

2 A Note on Notation 4

3 Heterogeneous Teams Technique 4

4 Hands-Off Technique 5

5 Incentive Contracts Technique 5

6 Heterogeneous Teams with Incentive Contracts Discussion 6

III Proving Propositions 1-4 and Lemma 1 8

7 Proving Propositions 1 and 4 8

8 Proving Proposition 2 9

9 Proving Proposition 3 10

10 Proof of Lemma 1 10

IV Further Discussion on Observations 13

11 On Observation 1 13

12 On Observation 2 14

13 On Observation 3 15

14 On Observation 4 15

V The Perfectly Aligned Agent and Other Extensions 16

15 Perfectly Aligned Agent 16
15.1 Full Equilibrium Strategy . . . 17
15.2 Proving Proposition 5 . . . 17
15.3 Foreign-Domestic Team or Perfectly Aligned Agent-Domestic Team? . . . 19

16 Averaged Actions: Proving Proposition 6 21

17 Agents Maximize Joint Utility 22
17.1 Proving Proposition 7 . . . 22

18 Expanding the Agent's Action Sets 23
18.1 Equilibrium Behavior . . . 23
18.2 Proving Proposition 7 . . . 24
18.3 Partial Comparative Statics on χ_d . . . 26
18.4 Discussion On Overshading . . . 27

19 Raising the Reservation Utility 27

Part I

Substantive Importance

1 Groups utilizing radical Islamist Foreign Fighters

To speak to the importance of the radical Islamist foreign fighter phenomenon, I did the following. I took the UCDP/PRIO battle deaths dataset (Pettersson and Öberg, 2020) for conflicts between 2010-2019. For each conflict ID, I then searched English-language media sources on these conflicts for documentation of radical Islamist transnational fighters joining radical, Islamist insurgent groups. If I could identify that these transnational fighters were not just coming from proximate regions, like from across a border,[1] I classified these conflict ID's as having foreign fighters. In the interest of full transparency, I include a table below that discusses each conflict ID that I coded as a "1", and a source justifying the coding.

To give some examples of what this procedure excluded (and why this should not be viewed as a holistic list of insurgencies that use foreign fighters or Islamist foreign fighters),[2] I did not document the following cases as instances that are applicable to my theory of radical Islamist foreign fighters:

- Several potentially radical Chechen fighters joined forces with the Ukranian government to fight in the Donbass, and the FARC in Colombia pulled in fighters from across Latin America and Europe. These are excluded because the foreign fighters were either themselves not Islamists, the foreign fighters were not joining a radical Islamist insurgent group, or both.
- Islamic State factions in Mozambique and Bangladesh currently have no documented cases of foreign fighters, and I found no evidence of foreign fighter assistance to Hamas of the PIJ during the time frame. These are excluded because there is no evidence of foreign fighter use, even though IS franchise groups commonly use foreign fighters.
- The OLF (fighting the government of Ethiopia) and various groups in the Central African Republic have used radical Islamist fighters from other states, but, best I could tell, these fighters came from contiguous states and I could not easily rule out that these fighters were truly "foreign." I will admit, by classifying these entries as "0", I may be under-reporting.

[1] As I characterize foreign fighters, these individuals do not have clear connections to a domestic population and are more ideologically motivated. The idea here is that a Saudi coming to fight for the Taliban would qualify as a foreign fighter, but a Pashto living on the Pakistan side of the Afghan-Pakistan border would not.

[2] For alternate datatsets, see Malet (2007) and Hegghammer (2010).

- While sources like Hegghammer (2010) describe foreign fighters being used in Kashmir, I do not find evidence that the foreign fighters joining Kashmiri groups were not from contiguous states (all evidence I found suggests they are from Afghanistan and Pakistan). This, in combination with work describing that that foreign fighters have never actually become a part of the Kashmiri independence movement led me to code this as a "0" (Siyech, 2018).

As two possible cases of over-identification, all conflict IDs involving Boko Haram and IS operating in Syria and Iraq I code as “1.” I do this despite, for example, there not being direct evidence of out-group radical Islamist foreign fighters in Boko Haram conducting operations against the Mali government (there is evidence of them in use against the Nigerian government). The treatment of foreign fighters from non-contiguous states as out-group members presents additional concerns with Boko Haram; because the Sahel region has a history of economic and social connectivity (Harmon, 2014), it is difficult to say that individuals travelling to fight with Boko Haram lack connections to the local population, even if they come from non-contiguous states. So to summarize, I may be over-identifying conflict IDs featuring Boko Haram (who eventually became IS) in the Sahel as falling within the scope of my theory.

The conflict ID's that I have listed here fall within the most direct scope of the theory. Of course, this is not the full set of cases of self-managing teams, but rather should be viewed as a list of cases most applicable to what I discuss in the paper.

Conflict id	Location	Side b	Start	End	Citation
230	Yemen (North Yemen)	AQAP, Ansarallah, Forces of Hadi	2010	2019	(Horton, 2017)
259	Iraq	Ansar al-Islam, IS	2010	2019	(Basit, 2014)
297	Nigeria	Jama'atu Ahlis Sunna Lidda'awati wal-Jihad	2011	2019	(Project, 2020)
299	Syria	Syrian insurgents	2011	2019	(Bakker *et al.*, 2014)
308	Philippines	ASG, MILF, BIFF, MNLF - NM, Maute group, al-Harakat al-Islamiyah, BIFF-K	2010	2019	(Beech and Gutierrez, 2019)
333	Afghanistan	Hizb-i Islami-yi Afghanistan, Taleban	2010	2019	(Giustozzi, 2019)
337	Somalia	Al-Shabaab, Hizbul Islam	2010	2019	(BBC, 2017)
353	Cameroon	Jama'atu Ahlis Sunna Lidda'awati wal-Jihad	2015	2017	(Zenn, 2018)
386	Algeria	AQIM, MUJAO	2010	2018	(Noonan, 2011)
395	Tajikistan	Forces of Mullo Abdullo, IMU	2010	2011	(ICG, 2011)
404	Pakistan	IMU, Jamaat-ul-Ahrar, Lashkar-e-Islam, TTP, TTP-TA	2010	2019	(Siddique, 2010)
432	Russia (Soviet Union)	Forces of the Caucasus Emirate	2010	2015	(Urban, 2015)
442	Mauritania	AQIM	2010	2011	(Boukhars, 2020)
11347	Mali	Ansar Dine, Military faction (Red Berets), AQIM, MUJAO, Signed-in-Blood Battalion, al-Murabitun, JNIM	2012	2019	(CEP, 2020a)
13588	Russia (Soviet Union)	IS	2015	2019	(Bakker *et al.*, 2014)
13604	Syria	IS	2013	2019	(Bakker *et al.*, 2014)
13637	Afghanistan	IS	2015	2019	(Youssef and Strobel, 2019)
13638	Cameroon	IS	2015	2019	(CEP, 2020b)
13639	Niger	IS	2015	2019	(ECF, 2019)
13640	Chad	IS	2015	2019	(CEP, 2020b)
13641	Nigeria	IS	2015	2019	(CEP, 2020b)
13645	Yemen (North Yemen)	IS	2015	2017	(Raghavan, 2019)
13646	Kenya	Al-Shabaab	2015	2019	(Scahill, 2015)
13648	Egypt	IS	2015	2019	(McManus, 2020)
13694	Libya	IS	2015	2019	(Anaizi, 2015)
13721	Algeria	IS	2015	2015	(Noonan, 2011)
13936	Pakistan	IS	2016	2019	(CSI, 2018)
14113	Mali	IS	2017	2019	(ECF, 2019)
14197	Somalia	IS	2019	2019	(Weiss, 2019)
14268	Iran	IS	2017	2017	(Basit, 2014)
14275	Philippines	IS	2016	2019	(Yusa, 2018)

Table 1: Insurgencies using out-group radical Islamist foreign fighters. All variables in the table follow UCDP/PRIO variable coding.

Part II

Full Equilibrium Strategies

2 A Note on Notation

Here I describe full equilibrium behavior within all techniques and I provide a more detailed discussion of the the Heterogeneous Teams with Incentive Contracts Technique.

Throughout the Appendix, I will be using a more detailed utility function than what was provided in the text. Whereas in the text agent i had utility function

$$U_i = \sum_{t=1}^{\infty} \delta^{t-1} \left(-|a_{i,t} - \chi_\tau| - |a_{j,t} - \chi_\tau| - \gamma |a_{i,t} - \omega_t| + G_{i,t} \right), \tag{1}$$

here agent i has utility function

$$U_i = \sum_{t=1}^{\infty} \delta^{t-1} \left(-\alpha |a_{i,t} - \chi_\tau| - \beta |a_{j,t} - \chi_\tau| - \gamma |a_{i,t} - \omega_t| + G_{i,t} \right), \tag{2}$$

where α and β are positive constants and $\alpha > \gamma$. Throughout this appendix, I will re-state the Propositions in the text in terms of the new utility function. So to summarize, Proposition 1 is equivalent to Proposition 1A (the equivalent appendix version) when $\alpha = \beta = 1$.

3 Heterogeneous Teams Technique

Using the appendix utility functions, I can state Proposition 1. First I will re-define the following:

Definition: $\tilde{z}_1$ and $\tilde{z}_2$ are defined as as

- $\tilde{z}_1 = 1$ and $\tilde{z}_2 = 1$ if $\tilde{k}_f \geq 1$,
- $\tilde{z}_1 = 1$ and $\tilde{z}_2 = \tilde{k}_f$ if $\tilde{k}_d \tilde{k}_f \geq 1$ and $\tilde{k}_f < 1$, and
- $\tilde{z}_1 = 0$ and $\tilde{z}_2 = 0$ if $\tilde{k}_d \tilde{k}_f < 1$,

where $\tilde{k}_d = \frac{\delta \beta \chi_f}{(\alpha - \gamma)(1 - \delta - \chi_d)}$ and $\tilde{k}_f = \frac{-\beta \delta \chi_d}{(\alpha - \gamma)(\chi_f + 1 - \delta)}$.

Next, I can state the Proposition 1A.

Proposition 1A: When the principal employs the Heterogeneous Teams Technique:

- Agents set $a_{1,t} = (1 - \tilde{z}_1)\chi_d + \tilde{z}_1\omega_t$ and $a_{2,t} = \tilde{z}_2\omega_t + (1 - \tilde{z}_2)\chi_f$ for all $t \in \{1, 2, 3, ...\}$,
- $\mathbb{E}U_p = ((1 - \tilde{z}_1)\chi_d - (1 - \tilde{z}_2)\chi_f) / (1 - \delta) - \kappa_o$,
- $\mathbb{E}U_1 = (\tilde{z}_1\chi_d - \beta((1 - \tilde{z}_2)\chi_f - \chi_d) - \gamma(1 - \tilde{z}_1)(-\chi_d)) / (1 - \delta)$,
- $\mathbb{E}U_2 = (-\alpha\tilde{z}_2\chi_f - \beta(\chi_f - (1 - \tilde{z}_1)\chi_d) - \gamma(1 - \tilde{z}_2)(\chi_f)) / (1 - \delta)$.

In the first stage, the principal sets $o_p = f$, $m = 0$, and $G_1 = G_2 = 0$. Also in this stage, both agents set $b_i = a$. In the second stage, in period $t = 1$, each agent i who is type τ selects action $a_{i,t} = \tilde{z}_i\omega_t + (1 - \tilde{z}_i)\chi_\tau$, with $\tilde{z}_1$ and $\tilde{z}_2$ defined in the text. For periods $t > 1$, if in period $t - 1$ agents select the actions characterized by $\tilde{z}_1$ and $\tilde{z}_2$, then in period t agent i selects the action characterized by $\tilde{z}_i$. For periods $t > 1$, if in period $t - 1$ either agent deviates from selecting the actions characterized by $\tilde{z}_1$ or $\tilde{z}_2$, then each agent i selects the actions characterized by $z_i = 0$ in period t and all future periods.

4 Hands-Off Technique

Using the appendix utility functions, I can state Proposition 2A.

Proposition 2A: When the principal employs the Hands-Off Technique:

- Agents set $o_a = d$, $a_{1,t} = \chi_d$, and $a_{2,t} = \chi_d$ ($z_1 = 0$ and $z_2 = 0$) for all $t \in \{1, 2, 3, ...\}$.
- $\mathbb{E}U_p = 2\chi_d/(1 - \delta)$,
- $\mathbb{E}U_1 = \mathbb{E}U_2 = \gamma\chi_d/(1 - \delta)$.

hy In the first stage, the principal sets $o_p = u$, $m = 0$, and $G_1 = G_2 = 0$. Also in this stage, agent 1 sets $o_a = d$, and both agents set $b_i = a$. In the second stage, both agents set $a_{i,t} = \chi_d$ for all t ($z_1 = z_2 = 0$).

5 Incentive Contracts Technique

Using the appendix utility functions, I can state Proposition 3A.

Proposition 3A: When the principal employs the Incentive Contracts Technique:

- Agents set $o_a = d$, $a_{1,t} = \omega_t$, and $a_{2,t} = \omega_t$ ($z_1 = 1$ and $z_2 = 1$)for all $t \in \{1, 2, 3, ...\}$.

- $\mathbb{E}U_p = \left(2\chi_d(\alpha-\gamma) - \kappa_m\right)/(1-\delta)$,
- $\mathbb{E}U_1 = \mathbb{E}U_2 = (\beta+\gamma)\chi_d/(1-\delta)$.

In the first stage, the principal sets $o_p = u$, $m = 1$, and $G_i(a_{i,t}) = (\alpha-\gamma)(a_{i,t} - \chi_d)$ for each agent i for all t. Also in this stage, Agent 1 sets $o_a = d$, and both agents set $b_i = a$. In the second stage, both agents set $a_{i,t} = \omega_t$ for all t ($z_1 = z_2 = 1$).

6 Heterogeneous Teams with Incentive Contracts Discussion

To summarize what occurs, in the first stage, the principal sets $o_p = f$, $m = 1$, and $G_{1,t}(a_1) = \hat{g}_1^* \left(a_{1,t,} - \chi_d\right)$ and $G_{2,t}(a_2) = \hat{g}_2^* \left(\chi_f - a_{2,t,}\right)$ for all t, where $\hat{g}_1^*$and $\hat{g}_2^*$ maximize the principal's expected utility from the agent's actions. I will refer to g_1 and g_2 as the "transfer constants." Also in this stage, both agents set $b_i = a$. In the second stage, in period $t = 1$, each agent i who is type τ selects action $a_{i,t} = \hat{z}_i\omega_t + (1-\hat{z}_i)\chi_\tau$, with $\hat{z}_1$ and $\hat{z}_2$ defined in the appendix. For periods $t > 1$, if in period $t-1$ agents select the actions characterized by $\hat{z}_1$ and $\hat{z}_2$, then in period t agent i selects the action characterized by $\hat{z}_i$. For periods $t > 1$, if in period $t-1$ either agent deviates from selecting the actions characterized by $\hat{z}_1$ or $\hat{z}_2$, then agent i selects the actions characterized by $z_i = 0$ in period t and all future periods.

So long that $g_1 < \alpha-\gamma$ and $g_2 < \alpha-\gamma$,[3] in equilibrium agents will select shading levels $\hat{z}_1$ and $\hat{z}_2$, which I introduce then describe below.

Definition: $\hat{z}_1$ and $\hat{z}_2$ are defined as

- $\hat{z}_1 = 1$ and $\hat{z}_2 = 1$ if $\hat{k}_d \geq 1$ and $\hat{k}_f \geq 1$,
- $\hat{z}_1 = 1$ and $\hat{z}_2 = \hat{k}_f$ if $\hat{k}_d\hat{k}_f \geq 1$ and $\hat{k}_f < 1$,
- $\hat{z}_1 = \hat{k}_d$ and $\hat{z}_2 = 1$ if $\hat{k}_d\hat{k}_f \geq 1$ and $\hat{k}_d < 1$,
- $\hat{z}_1 = 0$ and $\hat{z}_2 = 0$ if $\hat{k}_d\hat{k}_f < 1$.

with $\hat{k}_d = \frac{\delta\beta\chi_f}{(\alpha-\gamma-g_1)(1-\delta-\chi_d)}$ and $\hat{k}_f = \frac{-\beta\delta\chi_d}{(\alpha-\gamma-g_2)(\chi_f+1-\delta)}$.

Given these actions, I modify expression (2) to define the set of transfer constants $(\hat{g}_1^*, \hat{g}_2^*)$

[3] It is straightforward to show that the principal would never want to make offers $g_1 \geq \alpha-\gamma$ or $g_2 \geq \alpha-\gamma$.

that the principal will select from:

$$(\hat{g}_1^*, \hat{g}_2^*) \in \underset{g_1 \geq 0,\, g_2 \geq 0}{arg\,max} \left\{ \left((1 - \hat{z}_1(g_1, g_2) + \hat{z}_1(g_1, g_2) g_1) \chi_d - (1 - \hat{z}_2(g_1, g_2) + \hat{z}_2(g_1, g_2) g_2) \chi_f \right) / (1 - \delta) \right\}. \tag{3}$$

Because the principal's optimization function is neither continuous nor optimized over a closed interval, a natural concern is that under certain parameters a maximum does not exist. However, it does.

Lemma 1. The set of $(\hat{g}_1^*, \hat{g}_2^*)$ satisfying (3) is nonempty and satisfies $g_1 \leq \alpha - \gamma$ and $g_2 \leq \alpha - \gamma$.

Proof: See Section 10.

With Lemma 1 in place, the principal's and agent's actions can be described.

Proposition 4: When the principal employs the Heterogeneous Teams with Incentive Contracts Technique):

- Agents set $a_{1,t} = (1 - \hat{z}_1(\hat{g}_1^*, \hat{g}_2^*))\chi_d + \hat{z}_1(\hat{g}_1^*, \hat{g}_2^*)\omega_t$, and $a_{2,t} = \hat{z}_2(\hat{g}_1^*, \hat{g}_2^*)\omega_t + (1 - \hat{z}_2(\hat{g}_1^*, \hat{g}_2^*))\chi_f$ for all $t \in \{1, 2, 3, ...\}$.
- $\mathbb{E}U_p = \left((1 - \hat{z}_1(g_1^*, g_2^*) + \hat{z}_1(g_1^*, g_2^*) g_1^*) \chi_d - (1 - \hat{z}_2(g_1^*, g_2^*) + \hat{z}_2(g_1^*, g_2^*) g_2^*) \chi_f - \zeta \right) / (1 - \delta) - \kappa$,
- $\mathbb{E}U_1 = \left(\alpha \hat{z}_1(g_1^*, g_2^*) \chi_d - \beta \left((1 - \hat{z}_2(g_1^*, g_2^*)) \chi_f - \chi_d \right) - \gamma \left((1 - \hat{z}_1(g_1^*, g_2^*))(\omega_t - \chi_d) \right) - \hat{g}_1^* \chi_d \right) / (1 - \delta)$,
- $\mathbb{E}U_2 = \left(-\alpha \hat{z}_2 \chi_f - \beta \left(\chi_f - (1 - \hat{z}_1(g_1^*, g_2^*)) \chi_d \right) - \gamma \left((1 - \hat{z}_2(g_1^*, g_2^*))(\chi_f - \omega_t) \right) + \hat{g}_2^* \chi_f \right) / (1 - \delta)$.

Proof: See Section III

A simple example can demonstrate that this Technique can give the principal greater utility than only using Incentive Contracts. Consider a case where $\chi_f = 1$ and $\chi_d = -1$, $\delta = 1$ and $\beta = 0.5$. Under the incentive contracts technique, the principal must offer, in expectation, $\alpha - \gamma$ per-period. If the principal formed mixed teams and also used incentive contracts, a expected transfer value of $\alpha - \gamma - 0.4$ per-period would compel agents to match their actions to the state of the world, which is a clearly smaller transfer value. Under these parameter values, for a low enough κ, this technique can outperform Incentive Contracts. Empirically, here the principal brings in a diverse range of agents and would have weakly less subversion than the

Heterogeneous Teams Technique.

Part III

Proving Propositions 1-4 and Lemma 1

7 Proving Propositions 1 and 4

Because Proposition 1 follows from the case of Proposition 4 when $g_1 = g_2 = 0$, I prove these simultaneously. Based on Assumption 1, in equilibrium, agents shade by $z_1 \in [0,1]$ and $z_2 \in [0,1]$, and deviations from the equilibrium path are met with the grim-trigger punishment phase of agents setting $a_{1,t} = \chi_d$ and $a_{2,t} = \chi_f$ for all t. Also by Assumption 1, Agents will select the largest degree of shading. I fix the principal's transfers at g_1 and g_2, assuming that $g_1 < \alpha - \gamma$ and $g_2 < \alpha - \gamma$.

To examine which equilibria can be sustained, I consider the cases when agents shade towards a state of the world that is furthest from their ideal point. These are the cases that present the greatest incentive for agents to defect. For agent 1 this is $\omega_t = 1$, and for agent 2 this is $\omega_t = -1$. I first define several values.

Agent 1's worst 1 period payoff ($\omega_t = 1$) for remaining on the equilibrium path is
$U_1^{ON,W} = -\alpha\left(z_1 + (1-z_1)\chi_d - \chi_d\right) - \beta\left(z_2 + (1-z_2)\chi_f - \chi_d\right) - \gamma\left(1 - (z_1 + (1-z_1)\chi_d)\right) + g_1\left(z_1 + (1-z_1)\chi_d - \chi_d\right)$,

Agent 1's expected per-period utility for remaining on the equilibrium path is
$U_1^{ON,EU} = -\alpha\left((1-z_1)\chi_d - \chi_d\right) - \beta\left((1-z_2)\chi_f - \chi_d\right) - \gamma\left(-(1-z_1)\chi_d\right) + g_1\left((1-z_1)\chi_d - \chi_d\right)$.

Agent 1's utility from an optimal deviation from $\omega_t = 1$ is
$U_1^{OFF,W} = -\alpha\left(\chi_d - \chi_d\right) - \beta\left(z_2 + (1-z_2)\chi_f - \chi_d\right) - \gamma\left(1 - \chi_d\right)$.

Agent 1's expected per-period utility from being in the Nash reversion punishment phase is
$U_1^{OFF,EU} = -\alpha\left(\chi_d - \chi_d\right) - \beta\left(\chi_f - \chi_d\right) - \gamma\left(-\chi_d\right)$.

For agent 1 to remain on the equilibrium path, it must be that

$$U_1^{ON,W} + \frac{\delta}{1-\delta} U_1^{ON,EU} \geq U_1^{OFF,W} + \frac{\delta}{1-\delta} U_1^{OFF,EU},$$

which can be simplified to

$$z_1 \leq \frac{z_2 \delta \beta \chi_f}{(\alpha - \gamma - g_1)(1 - \delta - \chi_d)}.$$

A similar expression can be identified on the limits of z_2, which comes from considering agent 2 facing an $\omega_t = -1$. This is

$$z_2 \leq \frac{-z_1 \beta \delta \chi_d}{(\alpha - \gamma - g_2)(\chi_f + 1 - \delta)}.$$

These expressions are used to produce $\tilde{z}_1$ and $\tilde{z}_2$ for the Heterogeneous Teams Technique, and $\hat{z}_1$ and $\hat{z}_2$ for the Heterogeneous Teams with Incentive Contracts Technique. It follows from the agent's utility functions and reservation utilities that agents will both select $b_i = a$.

There are two items to note here. First, as g_1 and g_2 approach $\alpha - \gamma$, the right hand side of both expressions become greater than 1, meaning that, due to Assumption 1, transfers close to $\alpha - \gamma$ will not induce additional shading; also, in Lemma 1 I show that the principal does strictly worse using transfer values close to $\alpha - \gamma$. Second, so long that $0 \leq g_i < \alpha - \gamma$, the z_1 and z_2 are always positive.

8 Proving Proposition 2

If agent 1 selected a foreign type agent, in the repeated second stage, agents would select the strategies defined in the Heterogeneous Teams Technique. Selecting into a heterogeneous team produces a lower expected utility for agent 1 than selecting a domestic type partner (comparing agent 1's utilities in Proposition 1 and Proposition 2).

It is straightforward to see that a team of domestic type agents without receiving transfers does best setting $a_{1,t} = a_{2,t} = \chi_d$, and that the utilities from these actions exceeds each agent's reservation utility (making $b = a$ equilibrium behavior).

9 Proving Proposition 3

Using the appendix utility functions, I can state Proposition 2.

Proposition 3: When the principal employs the Incentive Contracts Technique:

- Agents set $o_a = d$, $a_{1,t} = \omega_t$, and $a_{2,t} = \omega_t$ ($z_1 = 1$ and $z_2 = 1$)for all $t \in \{1, 2, 3, ...\}$.
- $\mathbb{E}U_p = (2\chi_d(\alpha - \gamma) - \kappa_m)/(1 - \delta)$,
- $\mathbb{E}U_1 = \mathbb{E}U_2 = (\beta + \gamma)\chi_d/(1 - \delta)$.

With the offered transfer schedule $G_i(a_{i,t}) = (\alpha - \gamma)(a_{i,t} - \chi_d)$ for both agents i, if agent 1 selected a foreign type agent, the foreign type agent would always set $a_{i,t} = \chi_f$. This is strictly worse for agent 1 than selecting a domestic type agent 2.

When agent 1 and agent 2 are domestic type agents and are offered transfers of $G_i(a_{i,t}) = (\alpha - \gamma)(a_{i,t} - \chi_d)$, they are indifferent over all actions $a_{i,t} \in [\chi_d, \omega_t]$ (put another way, they are indifferent all shading levels $z_i \in [0, 1]$), which makes any set of actions within that range an equilibrium. By the maximization criterion on Assumption 1, agents will select $z_1 = z_2 = 1$. It is straightforward to see that the utilities from setting $z_1 = z_2 = 1$ exceeds each agent's reservation utility (making $b = a$ equilibrium behavior).

10 Proof of Lemma 1

I proceed by cases. In Cases 1 and 2, I define a closed set of (g_1, g_2) and show that all transfer constants outside of the set are either infeasible or strictly worse for the principal than values inside the closed set. I can then address any discontinuities to the principal's optimization function with the domain of the defined closed set, and I can show that in all cases a maximum still exists. In Case 3, I show that when the set I defined in the first case is empty, a unique maximum exists.

Case 1: $\frac{-\beta^2\delta^2\chi_d\chi_f}{(\alpha-\gamma)^2(1-\delta-\chi_d)(\chi_f+1-\delta)} < 1$

I define the set

$$\mathcal{G} = \begin{Bmatrix} (g_1, g_2) \, : \, g_1 \geq 0, \, g_2 \geq 0, \, g_2 \leq \alpha - \gamma + \dfrac{\delta^2\beta^2\chi_f\chi_d}{(\alpha - \gamma)(1 - \delta - \chi_d)(\chi_f + 1 - \delta)} \\ g_1 \leq \alpha - \gamma + \dfrac{\delta^2\beta^2\chi_f\chi_d}{(\alpha - \gamma)(1 - \delta - \chi_d)(\chi_f + 1 - \delta)} \end{Bmatrix} \tag{4}$$

which, by the Assumption of the case, is nonempty. Throughout the proof, I use values

$$g'_1 = \alpha - \gamma + \frac{\delta^2\beta^2\chi_f\chi_d}{(\alpha-\gamma)(1-\delta-\chi_d)(\chi_f+1-\delta)}$$

and

$$g'_2 = \alpha - \gamma + \frac{\delta^2\beta^2\chi_f\chi_d}{(\alpha-\gamma)(1-\delta-\chi_d)(\chi_f+1-\delta)}$$

where, by construction, $(g'_1, g'_2) \in \mathcal{G}$. As defined, g'_1 is a useful value because when the principal sets $G_{1,t}(a_1) = g'_1 * (a_{1,t,} - \chi_d)$ and $G_{2,t}(a_2) = 0$, then at these transfer values $\hat{k}_d * \hat{k}_f \geq 1$. Thus, any payment to Agent 1 greater than g'_1 is over-paying because it will not change the agents' actions. A similar logic holds for $g_2 = g'_2$ and $g_1 = 0$.

To show that values of (g_1, g_2) that fall outside of $\mathcal{G}$ are strictly worse for the principal requires a fairly tedious discussion of multiple cases. Before getting into the necessary casework, I introduce some notation. I define these transfer value pairs as $(\bar{g}_1, \bar{g}_2)$. I will abuse notation and let $\chi_d = \chi_1$ and $\chi_f = \chi_2$ as, within this case, agent 1 is domestic and agent 2 is foreign. Also, throughout this section, I define $i, j \in \{1, 2\}$, where $i \neq j$. Before proceeding, one final note – were it not for Assumption 1 (limiting to shading equilibria), there (a) would be open set issues where agents tries to select the largest or smallest action in an unbounded set, or (b) domestic agents may select shading levels larger than ω_t and foreign agents may select actions smaller than ω_t. In both cases, relaxing Assumption 1 would modify the process of the proof, but not the results.

When $\bar{g}_i \geq \alpha - \gamma$ and $\bar{g}_j \geq \alpha - \gamma$, the principal's transfers will induce agents to set agents set $a_{i,t} = a_{i,j} = \omega_t$ for all t. At transfer values g'_i and g'_j, agents set $a_{i,t} = a_{i,j} = \omega_t$ for all t (equivalent actions) at a transfer rate that, by definition, is less than that defined in $(\bar{g}_1, \bar{g}_2)$.

When $\bar{g}_i > \alpha - \gamma$ and $\bar{g}_j \in [0, \alpha - \gamma)$, then the principal's transfers induce agent i to select $a_{i,t} = \omega_t$ and will eliminate agent i's ability to use the Nash reversion punishment,[4] which results in agent j setting $a_{j,t} = \chi_j$ At transfer values g'_i and $\bar{g}_j$, agent i will select $a_{i,t} = \omega_t$ and agent j will shade some degree $0 \leq \hat{z}_j \leq 1$ (weakly more favorable actions) at a transfer rate that, by definition, is less than that defined in $(\bar{g}_1, \bar{g}_2)$.

When $\bar{g}_i \in (g'_i, \alpha - \gamma]$ and $\bar{g}_j \in [0, \alpha - \gamma)$, then the principal's transfers induce agent i to select $a_{i,t} = \omega_t$ while still allowing agent i the possibility of the Nash reversion punishment,

[4]At these transfer values, it is no longer a Nash equilibrium to set $a_{i,t} = 0$.

which results in agent j selecting some shading level $0 \leq \hat{z}_j \leq 1$. At transfer values g'_i and $\bar{g}_j$, agent i will select $a_{i,t} = \omega_t$ and agent j will shade some degree $0 \leq \hat{z}_j \leq 1$ (equivalent actions) at a transfer rate that, by definition, is less than that defined in $(\bar{g}_1, \bar{g}_2)$.

The examples above cover all possible transfer values falling outside of $\mathcal{G}$.

Having shown that all points outside of $\mathcal{G}$ are strictly worse for the principal, the original optimization problem is equivalent to optimizing over the closed set

$$(\hat{g}_1^*, \hat{g}_2^*) \in \underset{g'_1 \geq g_1 \geq 0,\, g'_2 \geq g_2 \geq 0}{arg\,max} \left\{ \left((1 - \hat{z}_1(g_1, g_2) + \hat{z}_1(g_1, g_2) g_1) \chi_d - (1 - \hat{z}_2(g_1, g_2) + \hat{z}_2(g_1, g_2) g_2) \chi_f \right) / (1 - \delta) \right\}.$$

This function possesses one discontinuity at $\hat{k}_d * \hat{k}_f = 1$. At this value, agents jump from not shading to some degree of shading; because the principal provides transfers when agents shade, based on the selected g_1 and g_2, at $\hat{k}_d * \hat{k}_f = 1$ the function could increase or decrease at the discontinuity. The principal's expected utility increases when the jump from not paying transfers (because agents set $\hat{z}_1 = 0$ and $\hat{z}_2 = 0$, the principal does not pay transfers) to paying transfers is productive and decreases when it is more cost than it is worth. I denote the set G'' as all pairs (g_1, g_2) such that $\hat{k}_d(g''_1) * \hat{k}_f(g''_2) = 1$. There are three sub-cases to consider here. First, consider if for all $(g''_1, g''_2) \in G''$ $EU_P(g_1 = 0,\, g_2 = 0) \leq EU_P(g_1 = g''_1,\, g_2 = g''_2)$. Note that the principal's expected utility from $g_1 = 0$ and $g_2 = 0$ is the same as the principal's utility from any (g_1, g_2) where $g_1 \leq g''_1$ and $g_2 \leq g''_2$, with one inequality holding strictly. In the first sub-case, the principal's optimization is upper semi-continuous and therefore attains its maximum over a closed set. Second, consider if some $(g''_1, g''_2) \in G''$ have the property $EU_P(g_1 = 0,\, g_2 = 0) > EU_P(g_1 = g''_1,\, g_2 = g''_2)$. Here the function is not upper semi-continuous, but the principal can either (a) select the (g''_1, g''_2) pair that does attain the maximum or (b) select the (g'''_1, g'''_2) where $\hat{k}_d(g'''_1) * \hat{k}_f(g'''_2) > 1$ that attains the maximum. Third, consider if for all $(g''_1, g''_2) \in G''$ $EU_P(g_1 = 0,\, g_2 = 0) > EU_P(g_1 = g''_1,\, g_2 = g''_2)$. Here the function is not upper semi-continuous, but the principal can either (a) select $g_1 = 0$ and $g_2 = 0$ which attains the maximum or (b) select the (g'''_1, g'''_2) where $\hat{k}_d(g'''_1) * \hat{k}_f(g'''_2) > 1$ that attains the maximum.

Case 2: $\frac{-\beta^2\delta^2\chi_d\chi_f}{(\alpha-\gamma)^2(1-\delta-\chi_d)(\chi_f+1-\delta)} \geq 1$ **and** $\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(\chi_f+1-\delta)} < 1$

In this case, any transfer values $g_1 > 0$ and $g_2 > \alpha - \gamma + \frac{\beta\delta\chi_d}{(\chi_f+1-\delta)}$ are counterproductive. Thus the principal is optimizing a continuous function over a closed set, implying that a maximum exists.

Case 3: $\frac{-\beta^2\delta^2\chi_d\chi_f}{(\alpha-\gamma)^2(1-\delta-\chi_d)(\chi_f+1-\delta)} \geq 1$ **and** $\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(\chi_f+1-\delta)} \geq 1$

When these hold, agents both setting $a_{i,t} = \omega_t$ is supported as an equilibrium without transfers. Thus, a maximum exists at $g_1 = g_2 = 0$. □

Part IV

Further Discussion on Observations

11 On Observation 1

First, I include the full statement of Observation 1A.

Observation 1A: Within a heterogeneous team:

- within the region where $\tilde{k}_d\tilde{k}_f \geq 1$, the principal's expected utility is weakly decreasing in α, weakly increasing in β and γ, and weakly decreasing in χ_d and χ_f.
- the expression $\tilde{k}_d\tilde{k}_f$ is decreasing in α, increasing in β and γ, and decreasing in χ_d. If a change in α, β, γ, or χ_d induces a change from $\tilde{k}_d\tilde{k}_f < 1$ to $\tilde{k}_d\tilde{k}_f \geq 1$ (or from $\tilde{k}_d\tilde{k}_f \geq 1$ to $\tilde{k}_d\tilde{k}_f < 1$), then the principal's expected utility is strictly increasing (or strictly decreasing) in that variable.
- the expression $\tilde{k}_d\tilde{k}_f$ is increasing in χ_f. If a change in χ_f induces a change from $\tilde{k}_d\tilde{k}_f < 1$ to $\tilde{k}_d\tilde{k}_f \geq 1$ or from $\tilde{k}_d\tilde{k}_f \geq 1$ to $\tilde{k}_d\tilde{k}_f < 1$, the effects on the principal's utility are ambiguous.
- within the region where $\tilde{k}_d\tilde{k}_f < 1$, the principal's expected utility is unchanging in α, β, and γ, strictly increasing in χ_d, and strictly decreasing in χ_f.

Now I can discuss further comparative statics not covered in the body of the paper.

Observation 1 reveals that the more agents know of and care about the actions of their teammates (larger values of β), the Nash reversion punishment phase becomes worse for the agents, which in turn can support a greater range of shading equilibria that are productive for the principal. Consistent with this theoretical expectation, militant groups do seem care about raising agents' intra-organizational awareness, which is one interpretation of β. For example, the Daesh newsletter and twitter account often distributed information about group members' activities, from their provision of public goods to beheadings. While communications that raise intra-organizational awareness may be beneficial for reasons other than self-managing teams, this model provides a new explanation for why raising the salience of others' activities within an organization can lead to greater productivity.

When $\tilde{k}_d\tilde{k}_f < 1$, the principal's expected utility is strictly increasing in χ_d and decreasing in χ_f. Within this range, agents do not shade and match their actions to their ideal points, meaning increases or decreases in χ_d and χ_f have direct effects on the agent's behavior, which directly affects the principal's utilities. It is worthwhile mentioning that if the principal ever through that parameter values were such that $\tilde{k}_d\tilde{k}_f < 1$, the principal would never use the Self-Managing Teams Technique because the principal could do strictly better by not incurring the κ cost and selecting the Hands-Off technique.

The cutpoint $\tilde{k}_d * \tilde{k}_f = 1$ separates the regions where agents do not shade from the regions where agents do shade. When χ_d decreases, for example, from χ_d to χ_d' with $\chi_d > \chi_d'$, and this results in a change from $\tilde{k}_d * \tilde{k}_f < 1$ to $\tilde{k}_d * \tilde{k}_f \geq 1$, Agents change from setting $a_{1,t} = \chi_d$ and $a_{2,t} = \chi_f$ to $a_{1,t} = \omega_t$ and $a_{2,t} = \chi_f - \tilde{z}_2(\chi_f - \omega_t)$. This shift always implies that agents are now closer to matching the principal's ideal actions. However, when χ_f increases, for example, from χ_f to χ_f' with $\chi_f < \chi_f'$, and this results in a change from $\tilde{k}_d * \tilde{k}_f < 1$ to $\tilde{k}_d * \tilde{k}_f \geq 1$, Agents change from setting $a_{1,t} = \chi_d$ and $a_{2,t} = \chi_f$ to $a_{1,t} = \omega_t$ and $a_{2,t} = \chi_f' - \tilde{z}_2(\chi_f' - \omega_t)$. This shift can lead to worse outcomes for the principal because if χ_f' is sufficiently very large, the new action $\chi_f' - \tilde{z}_2(\chi_f' - \omega_t)$ can be further from the principal's ideal point than χ_f was.[5]

12 On Observation 2

Observation 2A: When the principal employs the Hands-Off Technique, the principal's expected utility does not change with α, β, or γ. The principal's expected utility is increasing in χ_d and is unchanging in χ_f.

The Hands-Off Technique results in some standard ally-principle type results. As long as $\alpha > \gamma$ (as assumed by Assumption 0), agents want to subvert. Thus, the closer χ_d is to the principal's expected most-preferred action ($\mathbb{E}(\omega_t) = 0$), the better the principal will do.

[5]For example, when $\delta = 0.9$, $\alpha = 1.5$, $\beta = 1$, $\gamma = 0.65$, $\chi_d = -1$ and $\chi_f = 1$, then $\tilde{k}_d * \tilde{k}_f = 0.927$, the agents will not shade and the principal will receive a per-period expected payoff of -2 from self-managing teams. However, if all other parameters remain the same and now $\chi_f = 6$, then $\tilde{k}_d * \tilde{k}_f = 1.00$, agent 2 shades by $\tilde{z}_2 = 0.174$, and the principal will receive per-period payoff -5.96.

13 On Observation 3

Observation 3A: When the principal employs the Incentive Contracting Technique, the principal's expected utility is strictly decreasing in α, unchanging in β, and strictly increasing in γ. The principal's expected utility is strictly increasing in χ_d and is unchanging in χ_f.

Like the Hands Off Technique, the Incentive Contracts Technique also results in ally principle results. As χ_d increases, α decreases and γ increases, which means the agents' preferences are closer to those of the principal, it is less costly to buy good behavior through utility transfers.

14 On Observation 4

Observation 4A. When the principal employs the Heterogeneous Teams with Incentive Contracts Technique, the principal's expected utility is weakly decreasing in α and weakly increasing in β and γ.

I show comparative statics for α. Using the structure of this proof, similar results can be shown for β and γ.

By Lemma 1, there exists some nonempty set of transfer constants $(\hat{g}_1^*, \hat{g}_2^*)$ that maximizes the principal's expected utility function within the Heterogeneous Teams with Incentive Contracts Technique. I denote $(\hat{g}_1^*(\alpha), \hat{g}_2^*(\alpha))$ for an optimal set of transfer constants under parameter α, and I consider two possible α parameters, $\bar{\alpha}$ and $\underline{\alpha}$, where $\bar{\alpha} > \underline{\alpha}$. I will show that, in all cases, $(\hat{g}_1^*(\bar{\alpha}), \hat{g}_2^*(\bar{\alpha}))$ generates a weakly lower expected utility than $(\hat{g}_1^*(\underline{\alpha}), \hat{g}_2^*(\underline{\alpha}))$. Across cases, the proof relies on $\hat{k}_d$ and $\hat{k}_f$ (and $\hat{k}_d\hat{k}_f$) being strictly decreasing in α and strictly increasing in $\hat{g}_1$ and $\hat{g}_2$, which follows from first order conditions.

First, consider the case where some $(\hat{g}_1^*(\bar{\alpha}), \hat{g}_2^*(\bar{\alpha}))$ leads to $\hat{k}_d(\bar{\alpha}, \hat{g}_1^*(\bar{\alpha})) * \hat{k}_f(\bar{\alpha}, \hat{g}_2^*(\bar{\alpha})) < 1$. The principal's expected utility here is $U_p(\bar{\alpha}, \hat{g}_1^*(\bar{\alpha}), \hat{g}_2^*(\bar{\alpha}))$. As the first subcase, consider when, for transfer values $\hat{g}_1 = 0$ and $\hat{g}_2 = 0$, $\hat{k}_d(\underline{\alpha}, 0) * \hat{k}_f(\underline{\alpha}, 0) < 1$. Because the agents are not shading under $\bar{\alpha}$, $U_p(\bar{\alpha}, \hat{g}_1^*(\bar{\alpha}), \hat{g}_2^*(\bar{\alpha})) = U_p(\underline{\alpha}, 0, 0)$. And, because the principal selects an optimal transfer constant from the set that includes $\hat{g}_1 = 0$ and $\hat{g}_2 = 0$, I can claim $U_p(\underline{\alpha}, 0, 0) \leq U_p(\underline{\alpha}, \hat{g}_1^*(\underline{\alpha}), \hat{g}_2^*(\underline{\alpha}))$. By transitivity, in this subcase $\underline{\alpha}$ generates a weakly greater utility for the principal. As the second subcase, consider when, for transfer values $\hat{g}_1 = 0$ and $\hat{g}_2 = 0$, $\hat{k}_d(\underline{\alpha}, 0) * \hat{k}_f(\underline{\alpha}, 0) \geq 1$. Here agents are shading and the principal is not incurring any costs from transfers, so $U_p(\bar{\alpha}, \hat{g}_1^*(\bar{\alpha}), \hat{g}_2^*(\bar{\alpha})) < U_p(\underline{\alpha}, 0, 0)$. And, because the principal selects an optimal transfer value from the set that includes $\hat{g}_1 = 0$ and $\hat{g}_2 = 0$,I can claim

$U_p(\underline{\alpha}, 0, 0) \leq U_p(\underline{\alpha}, \hat{g}_1^*(\underline{\alpha}), \hat{g}_2^*(\underline{\alpha}))$. By transitivity, in this subcase, $\underline{\alpha}$ generates a strictly greater utility for the principal.

Second, consider the case where some $(\hat{g}_1^*(\bar{\alpha}), \hat{g}_2^*(\bar{\alpha}))$ leads to $\hat{k}_d(\bar{\alpha}, \hat{g}_1^*(\bar{\alpha})) * \hat{k}_f(\bar{\alpha}, \hat{g}_2^*(\bar{\alpha})) \geq 1$. I can define g_1' and g_2' as the following:

$$g_1' = \begin{cases} g_1' \text{ such that } \hat{k}_d(\underline{\alpha}, g_1') = \hat{k}_d(\bar{\alpha}, \hat{g}_1^*(\bar{\alpha})) & \text{if } \hat{k}_d(\underline{\alpha},0) \leq \hat{k}_d(\bar{\alpha}, \hat{g}_1^*(\bar{\alpha})) \\ 0 & \textit{otherwise} \end{cases}$$

and

$$g_2' = \begin{cases} g_2' \text{ such that } \hat{k}_f(\underline{\alpha}, g_2') = \hat{k}_f(\bar{\alpha}, \hat{g}_2^*(\bar{\alpha})) & \text{if } \hat{k}_f(\underline{\alpha},0) \leq \hat{k}_f(\bar{\alpha}, \hat{g}_2^*(\bar{\alpha})) \\ 0 & \textit{otherwise} \end{cases}$$

where, because $\hat{k}_d$ and $\hat{k}_f$ are decreasing in α and increasing in transfer constants, it must be that $g_1' \leq \hat{g}_1^*(\bar{\alpha})$ and $g_2' \leq \hat{g}_2^*(\bar{\alpha})$. Thus, $U_p(\bar{\alpha}, \hat{g}_1^*(\bar{\alpha}), \hat{g}_2^*(\bar{\alpha})) \leq U_p(\underline{\alpha}, g_1', g_2')$. Because the principal selects an optimal transfer value from the set that includes $\hat{g}_1 = g_1'$ and $\hat{g}_2 = g_1'$,I can claim $U_p(\underline{\alpha}, g_1', g_2') \leq U_p(\underline{\alpha}, \hat{g}_1^*(\underline{\alpha}), \hat{g}_2^*(\underline{\alpha}))$. By transitivity, $\underline{\alpha}$ generates a weakly greater utility for the principal. □

Part V

The Perfectly Aligned Agent and Other Extensions

15 Perfectly Aligned Agent

Here I will assume the perfectly aligned agent has utility function

$$U_{pa} = \sum_{t=1}^{\infty} \delta^{t-1} \left(-(\alpha + \gamma)|a_{pa,t} - \omega_t| - \beta |a_{j,t} - \omega_t| \right).$$

With this utility function, I define the following:

Definition: $\check{z}_1$ and $\check{z}_{pa}$ are defined as

- $\check{z}_1 = 0$ and $\check{z}_{pa} = 0$ if $\check{k}_d \check{k}_{pa} < 1$ and
- $\check{z}_1 = 1$ and $\check{z}_{pa} = \frac{1}{\check{k}_d}$ if $\check{k}_d \check{k}_{pa} \geq 1$,

where $\check{k}_d = \frac{-\beta\delta\chi_d}{(\alpha-\gamma)(1-\chi_d-\delta)}$ and $\check{k}_{pa} = \frac{-\beta\chi_d\delta}{(\alpha+\gamma)(1-\delta-\chi_d)}$.

And with this, I can state Proposition 5A.

Proposition 5A: Assume the principal forms a heterogeneous team with one domestic and one perfectly aligned agent.

- Agents set $a_{1,t} = \check{z}_1\omega_t + (1-\check{z}_1)\chi_d$ and $a_{pa,t} = (1-\check{z}_{pa})\omega_t + \check{z}_{pa}\chi_d$ for all t,
- $\mathbb{E}U_p = \left((1-\check{z}_1)\chi_d - \check{z}_{pa}(\omega_t - \chi_d)\right)/(1-\delta) - \kappa_o$.

15.1 Full Equilibrium Strategy

In period $t = 1$, the domestic agent (agent 1) selects $a_{1,t} = \check{z}_1\omega_t + (1-\check{z}_1)\chi_d$ and the perfectly aligned agent selects $a_{pa,t} = (1-\check{z}_{pa})\omega_t + \check{z}_{pa}\chi_d$, with $\check{z}_1$ and $\check{z}_{pa}$ defined in the text. For periods $t > 1$, if in period $t-1$ agents select the actions characterized by $\check{z}_1$ and $\check{z}_{pa}$, then in period t the domestic or perfectly aligned agent selects the action characterized by $\check{z}_1$ or $\check{z}_{pa}$ (respectively). For periods $t > 1$, if in period $t-1$ either agent deviates from selecting the actions characterized by $\check{z}_1$ and $\check{z}_{pa}$, then the domestic or perfectly aligned agent selects the action characterized by $\check{z}_1 = 0$ or $\check{z}_{pa} = 0$ (respectively) in period t and all future periods.

15.2 Proving Proposition 5

In equilibrium, agents shade by $z_1 \in [0,1]$ and $z_{pa} \in [0,1]$, and deviations from the equilibrium path are met with the grim-trigger punishment phase of agents setting $a_{1,t} = \chi_d$ and $a_{pa,t} = \omega_t$ for all t. The modification to Assumption 1 no longer implies that agents will select the largest degree of shading; rather they will select the degree of shading that benefits the principal the most. If the perfectly aligned agent selects $z_{pa} = 0$, then this will not induce any additional shading by the domestic agent. However, it can be possible for the perfectly aligned agent to move closer to agent 1's ideal point (set $z_{pa} > 0$) to induce agent 1 to shade closer to the state of the world in such a way that will benefit the principal.

Redefining terms used earlier, Agent 1's worst 1 period payoff ($\omega_t = 1$) for remaining on the equilibrium path is
$U_1^{ON,W} = -\alpha\left(1-\chi_d-(1-z_1)(1-\chi_d)\right) - \beta\left((1-z_{pa}) + \chi_d(z_{pa}) - \chi_d\right) - \gamma\left((1-z_1)(1-\chi_d)\right)$,

Agent 1's expected per-period utility for remaining on the equilibrium path is
$U_1^{ON,EU} = -\alpha\left((1-z_1)\chi_d - \chi_d\right) - \beta\left(\chi_d z_{pa} - \chi_d\right) - \gamma\left(-(1-z_1)\chi_d\right).$

Agent 1's utility from an optimal deviation from $\omega_t = 1$ is
$U_1^{OFF,W} = -\beta\left((1-z_{pa}) + \chi_d(z_{pa}) - \chi_d\right) - \gamma\left(1-\chi_d\right).$

Agent 1's expected per-period utility from being in the Nash reversion punishment phase is
$U_1^{OFF,EU} = \beta\chi_d + \gamma\chi_d.$

For agent 1 to remain on the equilibrium path, it must be that

$$U_1^{ON,W} + \frac{\delta}{1-\delta}U_1^{ON,EU} \geq U_1^{OFF,W} + \frac{\delta}{1-\delta}U_1^{OFF,EU},$$

which can be simplified to

$$z_1 \leq z_{pa}\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(1-\chi_d-\delta)}.$$

A similar expression can be identified for the limits on z_{pa}, which comes when the perfectly aligned agent faces a realization of $\omega_t = 1$. This is the "worst-case" for the perfectly aligned agent because the equation for shading implies that any $z_{pa} > 0$ here will result in the largest move away from ω_t. Disregarding the terms associated with β in the first period (because these will cancel out), the perfectly aligned agent's worst 1 period payoff ($\omega_t = 1$) for remaining on the equilibrium path is
$U_{pa}^{ON,W} = (-\alpha-\gamma)\left(z_{pa}(1-\chi_d)\right),$

Agent 1's expected per-period utility for remaining on the equilibrium path is
$U_{pa}^{ON,EU} = (-\alpha-\gamma)\left(z_{pa}(-\chi_d)\right) + \beta\left(1-z_1\right)\chi_d.$

Agent 1's utility from an optimal deviation from $\omega_t = 1$ is
$U_{pa}^{OFF,W} = 0.$

Agent 1's expected per-period utility from being in the Nash reversion punishment phase is
$U_{pa}^{OFF,EU} = \beta\chi_d.$

For agent 1 to remain on the equilibrium path, it must be that

$$U_{pa}^{ON,W} + \frac{\delta}{1-\delta}U_{pa}^{ON,EU} \geq U_{pa}^{OFF,W} + \frac{\delta}{1-\delta}U_{pa}^{OFF,EU},$$

which implies the following must hold.

$$z_{pa} \leq z_1 \frac{-\beta\chi_d\delta}{(\alpha+\gamma)(1-\delta-\chi_d)}.$$

z_1 and z_{pa} are how far a domestic agent and the perfectly aligned are willing to shade. For reasons similar to those expressed in the discussion on Proposition 1, non-zero levels of shading are possible when $\frac{\beta^2\chi_d^2\delta^2}{(\alpha+\gamma)(\alpha-\gamma)(1-\delta-\chi_d)^2} \geq 1$. Can increasing ever z_{pa} be beneficial for the principal? Re-writing the principal's expected per-period utility in terms of z_{pa} yields

$$U_p = (1 - \frac{-z_{pa}\beta\delta\chi_d}{(\alpha-\gamma)(1-\chi_d-\delta)})(\chi_d) + z_{pa}\chi_d,$$

where taking first order conditions yields

$$\frac{\partial U_p}{\partial z_{pa}} = \chi_d \left(\frac{\beta\delta\chi_d}{(\alpha-\gamma)(1-\chi_d-\delta)} + 1 \right).$$

Thus, U_p is increasing in z_{pa} when $\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(1-\chi_d-\delta)} > 1$ holds. Note that in order for $\frac{\beta^2\chi_d^2\delta^2}{(\alpha+\gamma)(\alpha-\gamma)(1-\delta-\chi_d)^2} \geq 1$, it must be that $\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(1-\chi_d-\delta)} > 1$.

As a final note, when $\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(1-\chi_d-\delta)} > 1$ holds, the principal does better having z_1 increase until it reaches the point where $z_1 = 1$ (agent 1 is matching action to the state of the world). Because $z_1 = z_{pa}\check{k}_d$, the principal does best up to the point where $z_{pa} = 1/\check{k}_d$. But is the perfectly aligned agent willing to make this shift? When $z_1 = 1$, the perfectly aligned agent is willing to shade up to $z_{pa} = \check{k}_{pa}$. Under the condition that $\check{k}_d\check{k}_{pa} \geq 1$, $\check{k}_{pa} \geq 1/\check{k}_d$, implying the perfectly aligned is willing to shade up to $1/\check{k}_d$.

Therefore, I can express the equilibrium levels of shading in regards to the $\frac{\beta^2\chi_d^2\delta^2}{(\alpha+\gamma)(\alpha-\gamma)(1-\delta-\chi_d)^2}$ condition, and use the above to produce equilibrium shading levels $\check{z}_1$ and $\check{z}_{pa}$.

15.3 Foreign-Domestic Team or Perfectly Aligned Agent-Domestic Team?

Here I provide a more detailed discussion on when the principal would prefer the foreign-domestic team over the perfectly aligned agent-domestic team. For ease, I refer to the domestic-foreign agent team as the D-F team and the domestic-perfectly aligned agent team as the D-PA team. I compare expected per-period utilities.

When $\tilde{k}_f \geq 1$, then the D-F team are setting $\tilde{z}_1 = \tilde{z}_2 = 1$, which grants the principal a greater expected utility than anything the D-PA team does. When $\tilde{k}_d \tilde{k}_f < 1$, then the D-F team is setting $\tilde{z}_1 = \tilde{z}_2 = 0$, which implies, for ally principle type reasons, the principal can do strictly better using the D-PA team.

For parameters where $\tilde{k}_d \tilde{k}_f \geq 1$ and $\tilde{k}_f < 1$, then whether D-F teams or D-PA teams are better for the principal depends on whether one of two cases holds.

Case 1: $\frac{\beta^2 \chi_d^2 \delta^2}{(\alpha+\gamma)(\alpha-\gamma)(1-\delta-\chi_d)^2} < 1$

The D-F team is better for the principal when

$$-(1 - \tilde{k}_f)\chi_f \geq \chi_d,$$

which can be re-written as

$$\chi_f \left(\frac{-\beta \delta \chi_d}{(\alpha - \gamma)(\chi_f + 1 - \delta)} - 1 \right) \geq \chi_d.$$

To offer some intuition on this condition, this inequality can hold or break depending on χ_f. Logically, when the foreign type agent is very extreme (possessing a large χ_f), shading can still occur, but the foreign fighter's shading will not result in a selected action close to ω_t. For example, when $\alpha = 1$, $\beta = 0.8$, $\gamma = 0.7$, $\chi_d = -2$, $\delta = 0.9$ and $\chi_f = 5$, the principal's per-period expected utility from the D-F team is ≈ -0.29 (with $\tilde{z}_1 = 1$ and $\tilde{z}_2 \approx 0.94$) and the principal's expected utility from the D-PA team is -2 (with $\check{z}_1 = \check{z}_{pa} = 0$). However, keeping all parameters but χ_f the same, when $\chi_f = 10$, the principal has per-period expected utility from the D-F team is ≈ -5.24 (with $\tilde{z}_1 = 1$ and $\tilde{z}_2 \approx 0.48$) and the per-period expected utility from the D-PA team is still -2.

Case 2: $\frac{\beta^2 \chi_d^2 \delta^2}{(\alpha+\gamma)(\alpha-\gamma)(1-\delta-\chi_d)^2} \geq 1$

The D-F team is better for the principal when

$$-(1 - \tilde{k}_f)\chi_f \geq \frac{1}{\check{k}_d} \chi_d,$$

Which can be re-written as

$$\chi_f \left(\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(\chi_f+1-\delta)} - 1 \right) \geq \frac{(\alpha+\gamma)(1-\delta-\chi_d)}{-\beta\delta}.$$

Similar to the previous case, this inequality can hold or break depending on χ_f.

16 Averaged Actions: Proving Proposition 6

I assume shading takes the structure defined in the paper, and that deviations from the defined shading path are met with the grim-trigger punishment phase of agents setting $a_{1,t} = \chi_d$ and $a_{2,t} = \chi_f$ for all t. To examine which equilibria can be sustained, I consider the cases when agents shade towards a state of the world that is furthest from their ideal point. These are the cases that present the greatest incentive for agents to defect. For agent 1 this is $\omega_t = 1$, and for agent 2 this is $\omega_t = -1$. I first define several values.

Agent 1's worst 1 period payoff ($\omega_t = 1$) for remaining on the equilibrium path is
$U_1^{ON,W} = -2z_1 - 2 - \gamma + \gamma z_1$,

Agent 1's expected per-period utility for remaining on the equilibrium path is
$U_1^{ON,EU} = \frac{1}{2}\alpha z_2 + \frac{\gamma z_2}{4} - \frac{1}{2}\alpha z_1 + \frac{\gamma z_1}{4} - \frac{\gamma}{2} - 2.$

Agent 1's utility from an optimal deviation from $\omega_t = 1$ is
$U_1^{OFF,W} = -2 - \gamma.$

Agent 1's expected per-period utility from being in the Nash reversion punishment phase is
$U_1^{OFF,EU} = -2 - \frac{1}{2}\gamma.$

For agent 1 to remain on the equilibrium path, it must be that

$$U_1^{ON,W} + \frac{\delta}{1-\delta} U_1^{ON,EU} \geq U_1^{OFF,W} + \frac{\delta}{1-\delta} U_1^{OFF,EU},$$

which can be simplified to

$$z_1 \leq \frac{z_2\delta\,(2+\gamma)}{(4-3\delta)(2-\gamma)}.$$

Due to symmetry, I can also say

$$z_2 \leq \frac{z_1 \delta \left(2 + \gamma\right)}{(4 - 3\delta)(2 - \gamma)}.$$

These expressions are used to derive the $\bar{z}_1$ and $\bar{z}_2$ conditions.

17 Agents Maximize Joint Utility

17.1 Proving Proposition 7

First, I state Proposition 7A.

Proposition 7A: Assume that agents maximize their joint per-period utility:

- Within the Incentive Contracts Technique, for agents $i \in \{1, 2\}$ and $j \in \{1, 2\}$ with $i \neq j$, the Principal transfers $G_{i,t} = (\alpha - \gamma)(a_{i,t} - \chi_d) + \beta(a_{j,t} - \chi_d)$. Agents set $a_{i,t} = \omega_t$. The principal receives expected payoff $EU_p = (2\chi_d(\alpha + \beta - \gamma) - \kappa_m)/(1 - \delta)$.
- Within the Heterogeneous Teams Technique, if $1 \leq (\beta)/(\alpha - \gamma)$ agents select $a_{i,t} = \omega_t$ and the principal receives expected payoff $EU_p = -\kappa_o$; otherwise, agents select $a_{1,t} = \chi_d$ and $a_{2,t} = \chi_f$ and the principal receives expected payoff $EU_p = (\chi_d - \chi_f)/(1 - \delta) - \kappa_o$.

For a heterogeneous team under Assumption 1, for any shading to occur, the condition $\tilde{k}_d \tilde{k}_f \geq 1$ must hold (as described in Proposition 1A). For a heterogeneous team under the assumptions here, agents will match their action to the principal's ideal point when $1 \leq \frac{\beta}{\alpha - \gamma}$ holds, which is both easier to satisfy than $\tilde{k}_d \tilde{k}_f \geq 1$ and generates a more favorable degree of shading for the principal. Put another way: even if agents are disregarding what the principal wants, by maximizing their joint utility, they will, under a broader parameter set, do precisely what is best for the principal.

In contrast, changing from Assumption 1 to the assumption that agents maximize their joint utility makes the Incentive Contracts Technique worse for principal. Comparing Proposition 3A to Proposition 7A, here the principal must pay each agent i an additional $\beta(a_{j,t} - \chi_d)$ to get agents to match their actions to the state of the world. While a transfer of $(\alpha - \gamma)(a_{i,t} - \chi_d)$ will make agent i indifferent over any action $a_{i,t} \in [\chi_d, \omega_t]$, agent i can still benefit when their teammate selects action χ_d (relative to action ω_t). Here the additional $\beta(a_{j,t} - \chi_d)$ transfer is necessary to make the team of agents jointly indifferent over any action $a_{i,t} \in [\chi_d, \omega_t]$. Thus, overall, making this change to Assumption 1 makes Heterogeneous Teams weakly better and Incentive Contracts more costly.

And now I prove Proposition 7A.

By matching action to the state of the world, a team of domestic agents receives joint expected utility $2(\alpha+\beta)\chi_d$. My matching action to their ideal points, a team of domestic agents receives joint expected utility $2\gamma\chi_d$. therefore, to properly motivate agents to match actions to the state of the world, the principal must transfer $G_{i,t} = (\alpha-\gamma)(a_{i,t}-\chi_d) + \beta(a_{j,t}-\chi_d)$ to both agents, which combined is an expected per-period transfer of $2(\alpha+\beta-\gamma)\chi_d$.

By matching action to the state of the world, a team of one domestic and one foreign agent receives joint expected utility $(\alpha+\beta)(\chi_d-\chi_f)$. My matching action to their ideal points, a team of domestic agents receives joint expected utility $-\beta(\chi_f-\chi_d)-\beta(\chi_f-\chi_d)-\gamma\chi_f+\gamma\chi_d$. Through algebra, the condition $1 \leq \beta/(\alpha-\gamma)$ must hold for a diverse team to fully self-manage.

18 Expanding the Agent's Action Sets

18.1 Equilibrium Behavior

In equilibrium, allowing for overshading means that a team with a domestic and a foreign agent will select shading levels $\mathring{z}_1$ and $\mathring{z}_2$, as follows:

Definition: $\mathring{z}_1$ and $\mathring{z}_2$ are defined as

- $\mathring{z}_1 = \tilde{z}_1$, $\mathring{z}_2 = \tilde{z}_2$ if $\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)} \leq 1$ or $\tilde{k}_f \geq 1$,
- $\mathring{z}_1 = \tilde{z}_1$, $\mathring{z}_2 = \tilde{z}_2$ if $\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)} > 1$, $\tilde{k}_f < 1$, and $0 < \mathring{k}_d \leq 1$,
- $\mathring{z}_1 = \mathring{k}_d$, $\mathring{z}_2 = \mathring{k}_f$ if $\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)} > 1$, $\tilde{k}_f < 1$, and $1 < \mathring{k}_d < \frac{1}{\tilde{k}_f}$,
- $\mathring{z}_1 = \frac{1}{\tilde{k}_f}$, $\mathring{z}_2 = 1$ if $\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)} > 1$, $\tilde{k}_f < 1$, and $\mathring{k}_d \geq \frac{1}{\tilde{k}_f}$,
- $\mathring{z}_1 = \frac{1}{\tilde{k}_f}$, $\mathring{z}_2 = 1$ if $\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)} > 1$, $\tilde{k}_f < 1$,
  and $(\alpha-\gamma)(\chi_f+1-\delta)(\alpha+\gamma)(1-\delta-\chi_d) \leq -\beta^2\chi_f\chi_d\delta^2$,

where $\mathring{k}_d = \frac{2\gamma(\alpha-\gamma)(\chi_f+1-\delta)(1-\delta-\chi_d)}{(\alpha-\gamma)(\chi_f+1-\delta)(\alpha+\gamma)(1-\delta-\chi_d)+\beta^2\chi_f\chi_d\delta^2}$ and $\mathring{k}_f = \frac{-2\gamma\beta\delta\chi_d(1-\delta-\chi_d)}{(\alpha-\gamma)(\chi_f+1-\delta)(\alpha+\gamma)(1-\delta-\chi_d)+\beta^2\chi_f\chi_d\delta^2}$.

Agent 1 is willing to overshade to levels $z_1 \leq \frac{2\gamma}{\alpha+\gamma} + z_2\frac{\beta\chi_f\delta}{(\alpha+\gamma)(1-\delta-\chi_d)}$ (so long that, as defined, $z_1 > 1$), and agent 2 is willing to shade to levels $z_2 \leq z_1\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(\chi_f+1-\delta)}$.[6] Because each agent's willingness to shade is an increasing functions of their teammates level of shading, when the

[6]Solving these expressions for one another yields the $\mathring{k}_d$ and $\mathring{k}_f$ terms.

domestic agent selects $z_1 > 1$, it can induce the foreign agent to select an action that is closer to the principal's ideal point relative to setting $z_1 = 1$ to an extent that may outweigh the disutility the principal receives from $z_1 > 1$. Thus, selecting $z_1 > 1$ can follow from the maximization criterion in Assumption 1, and this occurs when $\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)} > 1$ holds.[7]

Allowing for overshading sometimes does not induce any change in behavior (the first two bullet points), while at other times can produce efficiency gains for the principal (the remaining bullet points) In the conditions described in the first bullet point, overshading is not productive for the principal. When $\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)} \leq 1$, the expression $-|a_{1,t}(z_1) - \omega_t| - |a_{2,t}(z_2) - \omega_t|$ is not maximized through overshading, and when $\tilde{k}_f \geq 1$, overshading is unnecessary because both agents are willing to always set $a_{i,t} = \omega_t$ when placed on a heterogeneous team. In the second bullet point, overshading would be productive ($\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)} > 1$ and $\tilde{k}_f < 1$), but no feasible level of overshading is possible ($\mathring{k}_d \leq 1$). In the third bullet point, overshading is productive and agent 1 is willing to overshade, but agent 1 is unwilling to overshade to the degree such that agent 2 will match their action to the state of the world ($\mathring{z}_1 = \mathring{k}_d < \frac{1}{\tilde{k}_f}$, which induces $\mathring{z}_2 = \mathring{k}_f < 1$). In the forth bullet point, overshading is productive, agent 1 is willing to overshade, to the point that agent 2 matches their actions to the state of the world ($\mathring{z}_1 = \frac{1}{\tilde{k}_f}$, which induces $\mathring{z}_2 = 1$). In the final bullet point, overshading is productive, and the final inequality implies that $\frac{-\beta^2\delta^2\chi_f\chi_d}{(\alpha+\gamma)(1-\delta-\chi_d)(\alpha-\gamma)(\chi_f+1-\delta)} \geq 1$; when this is the case, agent 1 will always be willing to overshade to the level where $\mathring{z}_1 = \frac{1}{\tilde{k}_f}$.

I define equilibrium behavior and the principal's payoffs in Proposition 7.

Proposition 7: Assume $z_i \geq 0$. Using the Heterogeneous Teams Technique,

- Agents set $a_{1,t} = \mathring{z}_1\omega_t + (1 - \mathring{z}_1)\chi_d$ and $a_{i,t} = \mathring{z}_2\omega_t + (1 - \mathring{z}_2)(\chi_f)$ for all t,
- $\mathbb{E}U_p = ((1 - \mathring{z}_1)\chi_d - (1 - \mathring{z}_2)\chi_f)/(1 - \delta) - \kappa$.

As an important follow-up to Proposition 7, next I show that increasing χ_d can result in worse outcomes for the principal. Also next, I include a discussion on shading equilibria. Overall, expanding the agent's action sets can make Heterogeneous Teams better for the principal.

18.2 Proving Proposition 7

For reasons described in Proposition 1, agent 2's willingness to shade is $z_2 \leq \frac{-z_1\beta\delta\chi_d}{(\alpha-\gamma)(\chi_f+1-\delta)}$. When agent 1 selects a shading level $z_1 > 1$ (overshading), removing the β term and shading

[7] This condition is derived in the Appendix and follows from taking first order conditions of the principal's utility function with respect to agent 1's level of shading.

associated with it in the first period,[8] agent 1's worst 1 period payoff ($\omega_t = 1$) for remaining on the equilibrium path is
$U_1^{ON,W} = -\alpha\,(1-\chi_d-(1-k_d)(1-\chi_d)) - \gamma\,((k_d-1)(1-\chi_d))$,

Agent 1's expected per-period utility for remaining on the equilibrium path is
$U_1^{ON,EU} = -\alpha\,((1-k_d)\chi_d-\chi_d) - \beta\,((1-k_f)\,\chi_f-\chi_d) - \gamma\,(-(k_d-1)\chi_d)$.

Agent 1's utility from an optimal deviation from $\omega_t = 1$ (after removing the β term and shading associated with it) is $U_1^{OFF,W} = -\gamma\,(1-\chi_d)$.

Agent 1's expected per-period utility from being in the Nash reversion punishment phase is
$U_1^{OFF,EU} = -\beta\,(\chi_f-\chi_d) - \gamma\,(-\chi_d)$.

For agent 1 to remain on the equilibrium path, it must be that

$$U_1^{ON,W} + \frac{\delta}{1-\delta}U_1^{ON,EU} \geq U_1^{OFF,W} + \frac{\delta}{1-\delta}U_1^{OFF,EU},$$

which can be simplified to

$$z_1 \leq \frac{2\gamma}{\alpha+\gamma} + z_2\frac{\beta\,(\chi_f)\,\delta}{(\alpha+\gamma)(1-\delta-\chi_d)}.$$

The question remains if agent 1 selecting actions $z_1 > 1$ is valuable for the principal. Within this case, with z_1 and z_2 defined as the conditions above holding with equality, the principal has expected per-period utility $-(1-z_2)\chi_f + (z_1-1)\chi_d$. Substituting in $z_2 = \frac{-z_1\beta\delta\chi_d}{(\alpha-\gamma)(\chi_f+1-\delta)}$ and taking first order conditions with respect to z_1, the principal benefits from agent 1 setting $z_1 > 1$ when

$$\chi_d\left(1-\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)}\right) > 0,$$

which informs the inequalities involving $\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)}$.

The question also remains how far agent 1 is willing to shade. Substituting $z_2 = z_1\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(\chi_f+1-\delta)}$

[8] Because this is the one-period deviation payoff, agent 1 receives the same payoff stemming from agent 2's actions whether or not agent 1 remains on the equilibrium path.

into the expression $z_1 = \frac{2\gamma}{\alpha+\gamma} + z_2 \frac{\beta(\chi_f)\delta}{(\alpha+\gamma)(1-\delta-\chi_d)}$[9] and solving for z_1 yields

$$z_1 = \frac{2\gamma(\alpha-\gamma)(\chi_f+1-\delta)(1-\delta-\chi_d)}{(\alpha-\gamma)(\chi_f+1-\delta)(\alpha+\gamma)(1-\delta-\chi_d)+\beta^2\chi_f\chi_d\delta^2},$$

and a comparable equation can be solved for z_2 which is

$$z_2 = \frac{-2\gamma\beta\delta\chi_d(1-\delta-\chi_d)}{(\alpha-\gamma)(\chi_f+1-\delta)(\alpha+\gamma)(1-\delta-\chi_d)+\beta^2\chi_f\chi_d\delta^2}.$$

There are two things to note about these conditions. First, because any level of shading $z_2 > 1$ becomes unproductive for the principal, agent 1 will not select a shading level beyond $z_1 = \frac{(\alpha-\gamma)(\chi_f+1-\delta)}{-\beta\delta\chi_d}$. Therefore, when $\frac{(\alpha-\gamma)(\chi_f+1-\delta)}{\beta\delta\chi_d} < \frac{2\gamma(\alpha-\gamma)(\chi_f+1-\delta)(1-\delta-\chi_d)}{(\alpha-\gamma)(\chi_f+1-\delta)(\alpha+\gamma)(1-\delta-\chi_d)+\beta^2\chi_f\chi_d\delta^2}$, agent 1 will only shade to $z_1 = \frac{(\alpha-\gamma)(\chi_f+1-\delta)}{\beta\delta\chi_d}$. Second, the denominator in z_1 and z_2 as defined above ($(\alpha-\gamma)(\chi_f+1-\delta)(\alpha+\gamma)(1-\delta-\chi_d)+\beta^2\chi_f\chi_d\delta^2$) is not necessarily positive or non-zero. However, when $(\alpha-\gamma)(\chi_f+1-\delta)(\alpha+\gamma)(1-\delta-\chi_d)+\beta^2\chi_f\chi_d\delta^2 \leq 0$, it implies that $\frac{-\beta\delta\chi_d}{(\alpha-\gamma)(\chi_f+1-\delta)} * \frac{\beta(\chi_f)\delta}{(\alpha+\gamma)(1-\delta-\chi_d)} \geq 1$, which implies that each agent is willing to shade at a level greater than that of their teammate; this implies that overshading is always possible.

This discussion informs the equilibrium cases in the paper.

18.3 Partial Comparative Statics on χ_d

Whenever agent 1 and agent 2 select $\tilde{z}_1 = 1$ and $\tilde{z}_2 \in (0, 1]$, the principal's expected utility is decreasing in χ_d. Does this hold for levels of overshading? The following case analysis relies on for any $z_1 > 1$ and $z_2 \in [0, 1]$, the principal's expected utility is $U_p = -(1-z_2)\chi_f + (z_1-1)\chi_d$.

When $z_1 = \frac{(\alpha-\gamma)(\chi_f+1-\delta)}{-\beta\delta\chi_d} = \frac{1}{\tilde{k}_f}$, z_1 is increasing in χ_d. This means as χ_d increases, agent 1 shades more, which results in a lower expected utility for the principal (because in this case z_2 is unchanging).

When $z_1 = \mathring{k}_d$ and $z_2 = \mathring{k}_f$, the effect of changing χ_d on the principal's utility is ambiguous. Taking first order conditions and re-arranging yields

$$\frac{\partial U_p(\mathring{k}_d, \mathring{k}_f)}{\partial \chi_d} = \frac{2\gamma((\chi_f-\delta+1)(\alpha-\gamma)-\delta\beta\chi_f)\left((\alpha^2-\gamma^2)(\chi_f-\delta+1)(\chi_d+\delta-1)^2-\delta^2\left(\beta^2\right)\chi_f\chi_d^2\right)}{\left((-\chi_f+\delta-1)(\chi_d+\delta-1)(\alpha^2-\gamma^2)+\delta^2\beta^2\chi_f\chi_d\right)^2} - 1.$$

[9] Readers might wonder why in proposition 1 I did not substitute the comparable terms into one another. In the heterogeneous teams with no overshading, because agent 1 only shaded up to 1 and because agent 2 would never select a non-zero level of shading if $\tilde{k}_d < 1$, the expression would not have been correct. Here because agent 1 is selecting a level of $\frac{(\alpha-\gamma)(\chi_f+1-\delta)}{\beta\delta\chi_d} \geq z_1 > 1$, actually solving for this expression is necessary.

When the right hand side of the expression is negative, than the principal's expected utility is decreasing in χ_d. Admittedly, this statement is fairly complex, and I am unable to simply it further. However, using specified parameters, I am unable to find a case where, when $\frac{\beta\delta\chi_f}{(\alpha-\gamma)(\chi_f+1-\delta)} > 1$, $\tilde{k}_f < 1$, and $1 < \mathring{k}_d < \frac{1}{\tilde{k}_f}$ hold, where the first order conditions are positive. For example, when $\alpha = 1$, $\beta = 0.7$, $\gamma = 0.5$, $\chi_d = -1$, $\delta = 0.9$ and $\chi_f = 30$, the first order conditions are approximately -0.23. Whenever the first order conditions are negative, it implies that increasing χ_d makes the principal worse off, showing that non-ally principle type results can remain in the equilibrium with overshading.

18.4 Discussion On Overshading

Empirically, it is difficult to know what to make of overshading equilibria. Overshading equilibrium have the undesirable feature where agent 1 selects an action that they dislike and that, as a first-order effect, is bad for the organization. While overshading is in aggregate beneficial for the principal (because of the strategic response it induces in agent 2), it is decidedly more complex. While "nudging" agents towards non-zero shading equilibrium with $z_i \in [0, 1]$ can be thought of as the principal encouraging agents to do what's best (or close to what's best) for the organization because other agents are doing the same, nudging agents towards overshading equilibria would require convincing agent 1 to undertake an action that they do not like and that does not immediately benefit the organization. While it is possible to imagine select cases where the necessary complex internal practices leading to overshading are possible, it is hard to imagine that this sort of overshading is commonplace.

19 Raising the Reservation Utility

So far the agents have always done better by joining the group and participating in operations. Now I consider the case where the agents' reservation utility is raised to $R_a = 0$, which implies that the principal must pay a flat transfer rate across techniques to get agents to participate. Proposition 8 shows how this matters to the principal's utility across the Hands Off, Heterogeneous Teams, and Incentive Contracts Techniques. I do not discuss the agents' actions, as these remain the same as they are in preceding sections. To summarize what follows, when the agents' reservation utility binds, sometimes the principal must offer larger transfer amounts to agents within the Heterogeneous Teams Technique relative to the Incentive Contracts Technique. However, because transfers in the Heterogeneous Teams Technique can be flat-rate transfers that are not conditioned on the agents' actions, the principal avoids the per-period ζ payment, which can make Heterogeneous Teams less expensive than Incentive Contracts.

Proposition 8: Assume $R_a = 0$. To keep the agents from leaving the terror group:

- Within Incentive Contracts, the Principal transfers $G_{1,t} = (\alpha - \gamma)(a_{1,t} - \chi_d) - (\beta + \gamma)\chi_d$ and $G_{2,t} = (\alpha - \gamma)(a_{2,t} - \chi_d) - (\beta + \gamma)\chi_d$ for all $t \in \{1, 2, 3, ...\}$, and has $\mathbb{E}U_p = (2\chi_d(\alpha + \beta) - \zeta) / (1 - \delta)$,
- Within Hands-Off, the Principal transfers $G_{1,t} = -\gamma$ and $G_{2,t} = -\gamma$ for all $t \in \{1, 2, 3, ...\}$, and has $\mathbb{E}U_p = 2(\chi_d - \gamma) / (1 - \delta)$,
- Within Heterogeneous Teams, the Principal transfers $G_{1,t} = -\alpha\tilde{z}_1\chi_d + \beta((1 - \tilde{z}_2)\chi_f - \chi_d) + \gamma(1 - \tilde{z}_1)(-\chi_d)$ and $G_{2,t} = \alpha\tilde{z}_2\chi_f + \beta(\chi_f - (1 - \tilde{z}_1)\chi_d) + \gamma(1 - \tilde{z}_2)(\chi_f)$ for all $t \in \{1, 2, 3, ...\}$, and has $\mathbb{E}U_p = (2(\chi_d - \gamma) - G_{1,t} - G_{2,t}) / (1 - \delta) - \kappa$.

Among the three techniques examined here, using the Hands-Off Technique requires the smallest level of transfers. When $z_1 = 1$ and $z_2 = 1$, Heterogeneous Teams requires a greater transfer amount than Incentive Contracts. However, when $\tilde{k}_d\tilde{k}_f \geq 1$ and $\tilde{k}_f < 1$, then sometimes Heterogeneous Teams requires a smaller expected per-period transfer. When $\tilde{k}_f < 1$, agent 2 is selecting an action that is closer to agent 2's ideal point, and therefore does not need to be compensated as much to match their reservation utility.

The key take-away from Proposition 8 is that even with a high reservation utility, the principal may still use self-managing teams. While the principal sometimes pays a larger expected per-period transfer in the Heterogeneous Teams Technique than in the Incentive Contracts Technique, the principal does not need to pay ζ each period, which can make Heterogeneous Teams overall cheaper. Ultimately, while different agents do not want to work together without being provided with greater compensation, paying out a greater compensation can be worth the costs.

References

(2018) Islamic state khorasan (is-k), *Center for Strategic & International Studies*.

(2019) The islamic state in the greater sahara (isgs), *European Council on Foreign Relations*.

Anaizi, F. L.-A. M., Nader; Dotolo (2015) Confronting isis in libya: The case for an expeditionary counterinsurgency, *Small Wars Journal*.

Bakker, E., Paulussen, C. and Entenmann, E. (2014) Returning jihadist foreign fighters: Challenges pertaining to threat assessment and governance of this pan-european problem, *Security and human rights*, **25**, 11–32.

Basit, A. (2014) Foreign fighters in iraq and syria - why so many?, *Counter Terrorist Trends and Analyses*, **6**, 4–8.

BBC (2017) Who are somalia's al-shabab?, *BBC*.

Beech, H. and Gutierrez, J. (2019) How isis is rising in the philippines as it dwindles in the middle east, *New York Times*.

Boukhars, A. (2020) Keeping terrorism at bay in mauritania, *Africa Center for Strategic Studies*.

CEP (2020a) Mali: Extremism and counter-extremism, *Counter Extremism Project*.

CEP (2020b) Nigeria: Extremism and counter-extremism.

Giustozzi, A. (2019) *The Taliban at War: 2001-2018*, Oxford University Press.

Harmon, S. A. (2014) *Terror and insurgency in the Sahara-Sahel region: corruption, contraband, jihad and the Mali war of 2012-2013*, Ashgate Publishing, Ltd.

Hegghammer, T. (2010) The Rise of Muslim Foreign Fighters: Islam and the Globalization of Jihad, *International Security*, **35**, 53–94.

Horton, M. (2017) Fighting the long war: The evolution of al-qa'ida in the arabian peninsula, *Combatting Terrorism Center Sentinel*, **10**.

ICG (2011) Tajikistan: The changing insurgent threats, *International Crisis Group*.

Malet, D. (2007) The foreign fighter project.

McManus, A. (2020) Isis in the sinai: A persistent threat for egypt, *Center for Global Policy*.

Noonan, M. (2011) The foreign fighters problem, recent trends and case studies: Selected essays, *Foreign Policy Research Institute*.

Pettersson, T. and Öberg, M. (2020) Organized violence, 1989–2019, *Journal of peace research*, **57**, 597–613.

Project, C. E. (2020) Nigeria: Extremism & counter-extremism.

Raghavan, S. (2019) With the isis caliphate defeated in syria, an islamist militant rivalry takes root in yemen, *The Washington Post*.

Scahill, J. (2015) The purge: How somalia's al shabaab turned against its own foreign fighters, *The Intercept*.

Siddique, Q. (2010) Tehrik-e-taliban pakistan: An attempt to deconstruct the umbrella organization and the reasons for its growth in pakistan's north-west, *Danish Institute for International Studies (DIIS)*.

Siyech, M. S. (2018) Why has the islamic state failed to grow in kashmir?, *Counter Terrorist Trends and Analyses*, **10**, 11–15.

Urban, M. (2015) How many russians are fighting in ukraine?, *BBC*.

Weiss, C. (2019) Reigniting the rivalry: The islamic state in somalia vs. al-shabaab, *CTC Sentinel*, **12**.

Youssef, N. A. and Strobel, W. P. (2019) U.s. fights an islamic state rise in afghanistan, *Wall Street Journal*.

Yusa, Z. (2018) Philippines: 100 foreign fighters joined isis in mindanao since the marawi battle, *The Defense Post*.

Zenn, J. (2018) Boko haram's senegalese foreign fighters: Cases, trends and implications, *Jamestown Foundation*.