I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

Building Reputation: Proxy Wars and Transnational Identities

Marion Mercier
Arthur Silve
Benjamin Tremblay-Auger

Building Reputation: Proxy Wars and Transnational Identities

Marion Mercier
DIAL, LEDa, CNRS, IRD, Université Paris-Dauphine, Université PSL and IZA

Arthur Silve
Université Laval

Benjamin Tremblay-Auger
Stanford Graduate School of Business

ABSTRACT

Building Reputation: Proxy Wars and Transnational Identities*

In the context of a global security framework that mitigates interstate conflicts, nation-states establish a reputation for resolve by supporting foreign insurrections. Our theoretical predictions indicate that states with a greater number of co-ethnic groups abroad are more inclined to endorse foreign uprisings, resulting in higher political inclusion for such groups. With a dataset comprising over 280,000 ethnic group × sponsor state × target state × year observations, our analysis corroborates the role of reputation-building through co-ethnicity narratives as a determinant of sponsor state engagement. We argue that this mechanism engenders political concessions within target states, ultimately benefiting groups associated with reputed potential sponsors.

Corresponding author:
Marion Mercier
Université PSL
75016 Paris
France
E-mail: marion.mercier@cnrs.fr

1 Introduction

Involvement of foreign countries is an increasingly common feature of civil wars. In particular, proxy wars where a state supports a foreign rebellion have become more frequent in the past decades, while direct international conflict has been on a declining trend (Mumford, 2013 and San-Akca, 2016).[1] This may be the result of a global security architecture that provides a large space for plausible deniability when states sponsor rebel groups instead of engaging in a direct conflict with another state.[2] San-Akca (2016), who studied systematically states sponsoring foreign rebellions on the basis of ethnicity, religion or political ideology, proposes that they do so for two main reasons: the tactical delegation of their foreign policy,[3] and an emotional or ideological attachment to them. We argue that these two motives are intricately connected through a mechanism of reputation-building.

We show that states can instrumentalize their attachment to groups abroad in support of their foreign policy. A state may sponsor a rebellion in a target country to build its reputation as a dangerous neighbor, one that it is best to placate or appease with significant policy concessions. Anticipating this, a country facing the threat of rebellion might be more inclined to make policy concessions, in exchange for the sponsor's neutrality – the sponsor may then not need to demonstrate its attachment to the rebel group. The sincerity of this attachment is even inconsequential: its mere plausibility is enough for the target state to prefer avoiding the threat. In essence, we propose that sponsor states use proxy wars to build a reputation for resolve.[4]

---

[1]For just a few notable examples, Kausch (2017) and Erstad (2018) note Iranian support of Hezbollah, Qatari support of the Muslim Brotherhood (Ikhwan) and U.S. and Russian support of the Kurds. Evidence also shows Sudan providing a safe haven for the Ugandan rebels of the Lord's Resistance Army (Rohner et al., 2013), and Venezuela offering sanctuary to the Colombian group FARC (Martínez, 2017). Involvement of Uganda and Rwanda in the so-called first and second Congo wars in DRC is documented by McKnight (2015) and König et al. (2017), while Berman & Lake (2019) collect nine detailed case studies of such proxy wars.

[2]As suggested by an extensive literature, for instance Ellington (2003), Schultz (2010), Kibbe (2011), Bale (2012), Maoz & San-Akca (2012), Marshall (2016), and Farrow (2018). Cormac & Aldrich (2018) suggest that the term 'plausible deniability' is problematic, as the sponsor's involvement is usually an open secret (as Maurer (2015) shows in the example of Russia's involvement in Ukraine between 2014 and 2022). What matters for the sponsor is not having to acknowledge the interference. This is consistent with the mechanism that we develop, where the sponsor state can scare target states off but does not face potential international opprobrium.

[3]A dense literature supports that view, both empirically and theoretically. See in particular Bapat (2007, 2012), Byman & Kreps (2010), Salehyan (2010), Salehyan et al. (2011, 2014), Padró i Miquel & Yared (2012), Lee (2018), Konyukhovskiy & Grigoriadis (2018), Berman & Lake (2019) and Tremblay-Auger (2019).

[4]States that feel threatened may also sponsor rebel groups in neighboring countries to create a "buffer zone" which keeps enemy influence, troops and short-range missiles distant. This strategy, known as *forward/extended deterrence*, has been allegedly used by Russia in Eastern Europe (Lanoszka, 2016), and by Iran throughout the Middle East, e.g., by supporting the Hezbollah (Ahmadian & Mohseni, 2019). Developing a reputation for supporting rebels when needed is instrumental in keeping enemies at bay; the

Walter (2006, 2009) and Clare & Danilovic (2010) also study reputation-building, focusing respectively on separatist and inter-state conflicts. They find that the number of issues with third parties is an important factor in the initiation and escalation of a conflict, which is difficult to explain outside of the frame of reputation-building.[5] In a similar fashion, we show that the number of foreign groups with whom a potential sponsor state could have an ideational attachment is determinant both for the sponsor's propensity to support a rebel group in a given target state, and for the target state's propensity to grant political concession to this group. Again, such predictions are hard to explain in the absence of a reputation-building mechanism.

Reputation-building is notoriously hard to identify in international relations. Dafoe et al. (2014, p. 372) review the literature and conclude that "studying these concepts in a scientific manner faces a number of challenges intrinsic to the study of beliefs and motives more generally. Beliefs and motives are not directly observable, are subject to psychological and strategic biases in their expression, and are theoretically complex and context specific; furthermore, their behavioral implications are subject to substantial selection effects. Perhaps as a result, a coherent conceptual framework for these phenomena has remained elusive." Writing a formal model of states building their reputation using proxy wars allows us to probe deeper into the assumptions that underlie the mechanism. Even if we do not observe the motives of sponsor states, the model yields two testable predictions that would be hard to explain if not for the mechanism.

The model also makes the inherent endogeneity of the mechanism explicit. Indeed, the sponsor state and the rebel group have a common interest in overstating the strength of their link. It could then be the case that a state wishing to build a reputation conjures a tie with a rebel group in a neighboring country, or that the pre-existence of a tie helps it build a reputation. This is why our empirical analysis focuses on co-ethnicity as a narrative that lends credibility to a sponsor state's threat. Co-ethnic links are likely the closest approximation to exogenous ties, and variation in the number of a sponsor's co-ethnic connections plausibly serves as an objective source of variation in a state's motivation to build its reputation.

Relying on a sample of nearly 7,000 triads of ethnic group × target state × potential sponsor state, over the 1946-2010 period, we provide empirical evidence in support of the two theoretical predictions of the model. In particular, we show that an ethnic group is more likely to receive rebellion sponsorship from a co-ethnic foreign state when the latter counts more co-ethnic neighbors – in other words, more neighbors where its reputation could matter. We also find that groups sharing an ethnic tie with a potential

---

reputation mechanism studied here is key to forward deterrence.

[5]Note that, based on an alternative data set, Sambanis et al. (2018) do not find support for reputation-building when states face several potential future separatist groups.

sponsor which counts more co-ethnic neighbors benefit, on average, from more political inclusion in the target states.

All results are obtained in an arguably conservative empirical setting, given the nature of the two dependent variables. Indeed, as far as sponsorship is concerned, only cases of support to groups involved in an active rebellion are observed, while foreign states can also sponsor groups that do not rebel in practice. Turning to political concessions, we focus on one specific and demanding dimension – political representation – while it might be the case that target states offer political concessions under alternative forms to dangerous foreign sponsors.

Moreover, our findings emerge from specifications which, thanks to the triad-level geometry of the data, allow a rich set of fixed effects that neutralize the most endogenous sources of variation. Lastly, our results are robust to a myriad of robustness tests, which lends further credence to the reputation-building mechanism as a driver of state support to foreign rebellions.

The remainder of the paper proceeds as follows. Section 2 formally investigates the mechanism of reputation-building using proxy wars and derives non-trivial predictions from it. Moreover, it examines the mechanism in the context of the Russian case. Section 3 discusses our empirical approach to bring the predictions to the data, describes the data we rely on, and presents our econometric specifications. Section 4 lays out the empirical results and Section 5 concludes.

2 Theoretical approach

We propose a model to show how certain states build a reputation that they may sponsor foreign rebellions unless placated with foreign policy concessions from other states. Mechanisms of reputation-building are notoriously difficult to identify in data because of incommunicable motives. Writing a formal model, stylized as it is, allows us to discuss in details its underlying assumptions and check their validity in this context. It also allows us to predict regularities in the data that would be hard to explain in the absence of a reputation-building mechanism. In particular, the model associates a larger number of ties to groups abroad with a higher likelihood to support these groups' efforts to fight their respective governments, and a better political representation of these groups in their respective countries.
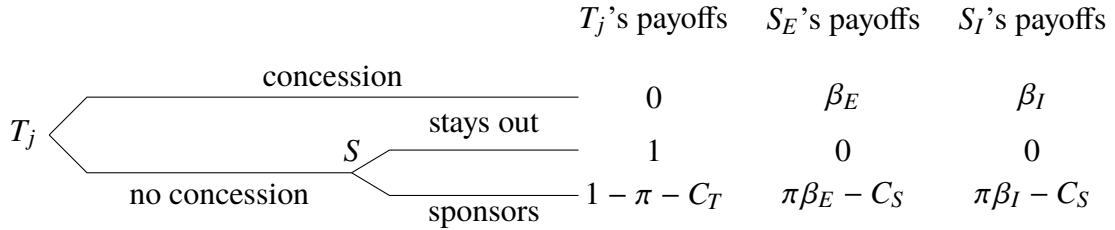
2.1 Formal model

Actors. A 'sponsor' state S successively faces, in n 'target' countries $j \in \{1, \ldots, n\}$ counting down from n to 1, the state T_j and a group that may rebel. State S wishes to

4

extract policy concessions from states T_j in negotiations over contentious regional or national issues, and may use asymmetric information about the strength of its connection with the rebel group as a negotiating tool.

Payoffs. In country j, the rebel group's utility is directly associated with its political representation / inclusion $x_j \in \{0, 1\}$, and the state's utility is $1 - x_j$. The sponsor's utility is $\sum_j \beta x_j$, with $\beta \geq 0$ a measure of the strength of the connection between the sponsor state and rebel groups.

Types. If the sponsor state cares strongly about the rebel groups' representation, $\beta = \beta_I$, the subscript I standing for an intrinsic motivation. If it cares less about the rebel group, $\beta = \beta_E$, with $0 < \beta_E < \beta_I$, E for an extrinsic motivation.[6]

Strategy sets and timing of the stage game. See the diagram below for an illustration. When S confronts T_j, T_j needs to take a decision regarding x_j. If $x_j = 1$, T_j confers political representation on the rebel group in country j: obviously, the sponsor state is satisfied, and therefore, the game reaches its next stage without foreign intervention. If $x_j = 0$, T_j refuses to accommodate the rebel group. In response, S decides whether or not to intervene and support the rebels. With foreign support, a rebellion occurs (Balch-Lindsay et al., 2008 and Bak et al., 2020), which imposes a cost C_S on the sponsor state S and C_T on state T_j.[7] The rebels win with probability $\pi \in (0, 1)$. If they win, they forcefully gain political representation ($x_j = 1$, with corresponding state payoffs $\beta - C_S$ and $-C_T$). If they lose, T_j does not need to make a concession to the rebels ($x_j = 0$, with corresponding state payoffs $-C_S$ and $1 - C_T$).

	T_j's payoffs	S_E's payoffs	S_I's payoffs
concession	0	β_E	β_I
stays out	1	0	0
sponsors	$1 - \pi - C_T$	$\pi\beta_E - C_S$	$\pi\beta_I - C_S$

Repetition. S faces each state T_j successively, counting down from n to 1. j is therefore an index of how many stage games remain to be played. Exposing the model

---

[6]The reputation literature usually characterizes the high-β sponsor as the 'commitment', 'Stackelberg', or 'behavioral' type, and the low-β as the 'normal' type. Dafoe et al., 2014 also call a low-β motivation 'instrumental.' We prefer the intrinsic / extrinsic terminology because it highlights that supporting a rebellion abroad may be an end in itself, or serve to build the sponsor's reputation.

[7]Certain states may also be concerned about neighboring civil wars having domestic repercussions. A conflict in its neighborhood increases the odds that a country will also experience conflict. Salehyan & Gleditsch (2006) find that the odds increase by about 52%, and Buhaug & Gleditsch (2008) by 44% (keeping the country-specific attributes usually emphasized in the comparative civil war literature constant). Braithwaite (2010) suggests that state capacity allows a state to withstand the threat of contagion, but also finds comparable estimates. It is no wonder, then, that states often help their neighbors *quell* rebellions, in order to avoid a contagion effect (Kathman, 2010, 2011) or more generally stave off the development of a regional conflict complex (Ansorg, 2011 and Silve & Verdier, 2018).

in terms of successive face-offs with n adversaries is simpler, although it would be more general to interpret n as a number of issues on which S expects it will need to negotiate with other states.[8]

Beliefs. Only S knows whether it is intrinsically or extrinsically motivated to intervene. In stage j, all other actors share a common belief that S is intrinsically motivated with probability p_j, and extrinsically with probability $1 - p_j$. This belief depends on the past actions of S, and in the first stage we take $p_n = \theta$.

Further assumptions. This general setup can lead to a variety of observations, some of them exceeding our ambition. To narrow down the discussion to the reputation-building mechanism, let us also assume that the expected payoff of supporting a rebellion is positive for an intrinsically motivated state, *i.e.* $\pi\beta_I - C_S > 0$, and negative for an extrinsically motivated one, *i.e.* $\pi\beta_E - C_S < 0$, and that a concession from another state is valuable enough that an extrinsically motivated state may be willing to pay the cost of supporting a rebellion, *i.e.* $(1 + \pi)\beta_E - C_S > 0$. We also assume that a state would prefer to avoid having to manage a rebellion, *i.e.* $1 - \pi - C_T < 0$.

Equilibrium. The setup parallels exactly that of Kreps & Wilson (1982), who studied the so-called 'chain-store' paradox. They demonstrate the existence of a unique sequential equilibrium to display 'plausible' beliefs (plausible beliefs p_j that the sponsor is intrinsically motivated are revised upwards when the sponsor fights). This equilibrium can be written as follows:

1. If the history of the play up to stage $j < n$ includes any instance that a state T_l ($l > j$) refused a concession yet state S did not retaliate by supporting a rebellion there, then the commonly shared belief that state S is driven by an intrinsic motive is $p_j = 0$. If S retaliated against every refusal so far, and the last refusal occurred in stage $m > j$, then $p_j = \max\{(\pi + C_T)^{1-m}, \theta\}$.

2. If S is intrinsically motivated to intervene, it always does.

---

[8]Such as it is exposed, the repetition of the game may generalize to any exogenous process by which n occasions of conflict occur with S's targets (e.g. a Poisson process). Fudenberg & Kreps (1987) show for instance that under some conditions, the sequential model is similar to a number of targets facing the sponsor simultaneously. As it happens, this means the number of targets is only an imperfect proxy of the motivation to build a reputation, since contentious issues are not necessarily distributed uniformly across targets and across sponsors. Yet, there are at least two ways in which we could extend the model. First, S may choose to establish its reputation on a given issue with a specific target, suggesting an endogenous process that we do not account for. Second, the occurrence of conflicts may be uncertain, suggesting an indefinitely repeated game. One difficulty with an infinitely or indefinitely repeated game is that the equilibrium is either separating or pooling – making reputations permanent and hard to find empirically. Mailath & Samuelson (2015) inventory the literature and find that impermanent reputations in infinitely repeated games may be the product of either of two assumptions: if the targets monitor the sponsor's actions imperfectly, or if the motivation of the sponsor changes randomly. Başer (2021) uses the latter in a model of reputation-building in international relations.

3. If S's motivation is extrinsic and the target state refuses to make a concession, it does not intervene when $j = 1$. At stage $j > 1$, if $p_j \geq (\pi + C_T)^{1-j}$, it intervenes, and if $p_j < (\pi + C_T)^{1-j}$, it intervenes with probability $((\pi + C_T)^{j-1} - 1)p_j/(1 - p_j)$, and does not with the complementary probability.

4. If $p_j > (\pi + C_T)^{-j}$, state T_j makes a concession. If $p_j < (\pi + C_T)^{-j}$, it refuses to do so. If $p_j = (\pi + C_T)^{-j}$, it randomizes, making a concession with probability $\beta_E/(C_S - \pi\beta_E)$.

2.2 Discussion of the equilibrium

This is a highly stylized, but rich setup. The game starts at $j = n$ and $p_j = \theta$ in one of four situations. a) For a large enough j, *i.e.* $j \geq 1 - \ln p_j/\ln(\pi + C_T)$, whatever the nature of its motivation, S would intervene; therefore T_j would concede rather than face a certain rebellion. b) If $1 - \ln p_j/\ln(\pi + C_T) > j > -\ln p_j/\ln(\pi + C_T)$, even if T_j is not certain S would intervene, the likelihood that it would not is too small for T_j to take the risk, and S still does not intervene. c) If $j = -\ln p_j/\ln(\pi + C_T)$, T_j randomizes, making a concession with probability $\beta_E/(C_S - \pi\beta_E)$. d) Finally, if $-\ln p_j/\ln(\pi + C_T) > j$, T_j denies the concession. In both situations c and d, an extrinsically motivated S_E would intervene with probability $((\pi + C_T)^{j-1} - 1)p_j/(1 - p_j)$ to maintain the ambiguity[9] and build its reputation (an intrinsically motivated S_I would of course intervene).

The four situations a, b, c, and d in which the game can start are also the four situations in which the game finds itself in each stage on the equilibrium path. To understand the equilibrium, it is useful to examine the sequence in which they occur (see diagram below). If and as long as there are enough remaining issues to negotiate, the situation remains in situation a. There comes a stage where the number of remaining issues is not large enough that an extrinsically motivated state would always intervene, but the threat is still enough to deter the target in that stage (situation b).[10] In the stage that follows, the target state tests the motivation of the sponsor (situation d). If the sponsor demonstrates resolve and retaliates, the following target state's beliefs make it indifferent between conceding or not: as long as the sponsor does not give in, and thus reveals its extrinsic motivation, the game remains in situation c. If in any stage, the state S does not retaliate against a refusal to concede, it reveals that its motivation is extrinsic, and expects no more concessions. In equilibrium, this may only happen in situations c

---

[9]van Houten (1998) offers another perspective on why a sponsoring state might wish to maintain ambiguity. If the state's support for the co-ethnic group is too forceful, it risks a direct confrontation with the target state. Conversely, if its support is perceived as insufficient, it exposes the co-ethnic group to potential repression. This delicate balance necessitates a strategy shrouded in ambiguity.

[10]In an unlikely scenario, the threat may make the target indifferent between conceding or not, *i.e.* switch directly to situation c.

or d. The game of reputation-building 'ends.' Finally, when $j = 1$, the state reveals the nature of its motivation, and the game 'ends' (not represented in the diagram below).
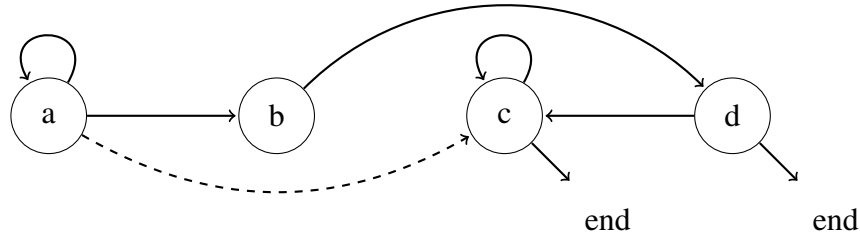
To summarize, here are the important intuitions of this equilibrium. A state S_I with an intrinsic motivation always supports a rebellion by a connected group abroad. Faced with such a sponsor, target states grant concessions to local rebels to prevent costly rebellions. A state S_E with an extrinsic motivation may also support a rebellion, even at a cost: this maintains an ambiguity about its motivation, and allows it to obtain concessions from other targets later on. It is all the more interesting for S_E to maintain an ambiguity that there is a larger number of issues on which S expects it will need to negotiate with other states.

Early in the game, if n is large enough, a state T_j expects that S would intervene, whatever the nature of its motivation. To avoid a certain confrontation, it prefers to make a concession. S's motivation remains untested during that phase. The common belief that it may have an extrinsic motivation remains constant. Later in the game (or in a game that starts with a smaller n), the set of concessions that S_E hopes to obtain shrinks. The likelihood that it would actually support a rebellion diminishes. Target states may start taking their chance, refusing a concession to the rebel group, hoping that S will choose not to intervene (thereby revealing its extrinsic motivation).

2.3   Implications of the model

We are particularly interested in the following two comparative static predictions. They are not trivial, and hard to interpret without the lens of a reputation-building mechanism. If supported by the data, these predictions would suggest that sponsors indeed support foreign rebellion with the view that it may help them extract concessions in future negotiations.

Result 1 *The probability that a state sponsors a rebellion in a given target country is non-monotonous in the number of targets in which it could sponsor other rebellions. It increases until the number is large enough that any target would rather concede than face a certain rebellion.*

Suppose we start at an arbitrary subgame with target T_j. Without loss of generality, we ignore past actions and write $j = n$, $T_j = T_n$ and $p_j = \theta$ as if we were in the first stage of a new game. Then, suppose that n increases for an exogenous reason (we discuss those in Section 3). For S to sponsor a rebellion in T_n, we first need T_n to refuse to concede, and then S to retaliate. In the equilibrium above, T_n would concede for n high enough (*i.e.* situations a and b). It would not concede in situation d. It would concede with a probability $\beta_E/(C_S - \pi\beta_E)$, independent of n, when n is low enough (*i.e.* in situation c). Meanwhile, in situations c and d, the probability that S retaliates is $(\pi + C_T)^{n-1}\theta$, increasing in n.[11]

Result 2 *The political representation of a rebel group increases in the number of targets in which a potential foreign sponsor could also sponsor rebellions.*

We need to consider three possible situations. If n is larger than $-\ln\theta/\ln(\pi + C_T)$ (*i.e.* in situations a and b), T_n simply concedes political representation to the rebel group. If n is smaller than $-\ln\theta/\ln(\pi + C_T)$ (*i.e.* in situations d), T_n does not concede, and S retaliates with probability $(\pi + C_T)^{n-1}\theta$ (increasing with n). Retaliation from S results in representation for the rebel group with probability π. T_n would concede with constant probability $\beta_E/(C_S - \pi\beta_E)$ in situation c, and if it does not, S would retaliate with probability $(\pi + C_T)^{n-1}\theta$ (still increasing in n).

Note that these two predictions are true *in all subgames*, meaning that an exogenous change in the number of targets (our key independent variable) has the same qualitative effect regardless of where we are in the sequence of the game. This is important since, in reality, reputation-building is a long-term dynamic process and changes in exogenous parameters may happen at any time.

2.4 Further remarks

In practice, foreign policy negotiations may take different forms than conceding representation to a rebellious group. Some concessions may even be less costly to the government of the target state than sharing power (cost smaller than 1), and simultaneously more advantageous to the sponsor state (payoff larger than β_I). Military cooperation, free trade agreements, mining concessions, for instance, are sometimes preferable to political representation of the co-ethnic group even for an intrinsically motivated sponsor (Fearon, 1995, Acemoglu, 2003, Powell, 2013, and Gates et al., 2016). Such a concession, preferable to both target and sponsor states, means that the equilibrium described above can also apply to the case where extrinsically motivated

---

[11]We are considering the probability of *any type* retaliating. With $p_j = \theta$, we get:
$(1 - \theta)((\pi + C_T)^{j-1} - 1)\frac{\theta}{1-\theta} + \theta \times 1 = (\pi + C_T)^{n-1}\theta.$

9

sponsors truly did not care about the political representation of their co-ethnic group, *i.e.*, if $\beta_E = 0$. The intuitions of the mechanism are unchanged in this case, and while the precise expressions to describe the mixed strategies and the beliefs are different, the comparative static predictions continue to apply (results available upon request).

In practice as well, the sponsor has a degree of discretion over what it would consider as a cause for intervention or retaliation, and over the initiation and escalation of disputes (Clare & Danilovic 2010, Barnhart 2021, and Başer 2021). It means that the sponsor needs not wait for the next occasion to build its reputation. It may even sometimes decide when the stage confrontation starts.

Although state sponsorship of rebel groups is usually relatively cheap, budget constraints may limit the ability of a sponsor to intervene again, once it has already intervened elsewhere. As the number of targets increases, budget constraints make it increasingly unlikely that a sponsor is able to retaliate against every challenge. Therefore, if anything, budget constraints would make our results 1 and 2 harder to verify empirically. The probability of support (and concession) would still rise with the number of target states, but less than if the budget constraint is slack.

More important, the sponsor sets the narrative of which rebel groups it is liable to support, and of what countries are its targets. To which extent can we take such narratives and declarations at face value? The rebel group and the sponsor state may have a common interest in overstating the strength of their tie (regardless of its nature), or even make it up to give credibility to the sponsor state's threat to support the rebellion (San-Akca, 2016). At the limit, that tie could even be immaterial except as a narrative that gives credibility to the threat. This creates a difficulty, as ties that matter for the mechanism we are describing may be endogenous to the mechanism itself.

2.5   Understanding Russia's imperialism in the "Near Abroad"

The Putin era (2000-) has been marked by a combination of direct intervention and indirect support from Russia in neighboring countries, even before the support for separatist groups in the Donbas region and the 2022 war in Ukraine (Allison, 2013).[12] Nevertheless, the nature of Russia's motivation remains ambiguous. Although it claims to have an intrinsic motivation to defend its Russian 'compatriots' in the near abroad, its actions often suggest more state-centric concerns.

The Compatriots doctrine has been a central element of Russia's foreign policy

---

[12]In the subsequent empirical section, we concentrate on indirect support for consistency, and because it allows us to test our predictions on significantly more observations than if we focused on direct intervention. However, the theoretical mechanism we have described could manifest in other forms, such as direct intervention. This is particularly relevant for a country that perceives itself as a great power (Rich, 2009 and Smith, 2014).

towards its neighboring countries since the dissolution of the USSR (Grigas, 2016). This narrative not only justifies Russia's interventions, it empowers Russian-speakers in target countries, thus influencing their domestic policies and regional dynamics (Batta, 2021). Intriguingly, its inconsistent use over time and across potential target countries suggests that state-centric motivations may hold greater importance than genuine concern for Russian-speaking minorities, with the protection issue conspicuously resurfacing when Russia's state interests are directly at stake (Simonsen, 2001, Pigman, 2019, Pieper, 2020, and Bendix, 2022).

Our theory offers suggestive insights into the puzzles raised by Russia's interactions with the independent Republics that emerged from the dissolution of the USSR.

Why is Russia so deeply committed to the 'Compatriots' narrative (Ziegler, 2006, Grigas, 2016, and Pieper, 2020)? With a significant number of ethnic links to Russian minorities (large N) across 14 former Soviet Republics (including 8 direct neighbors), Russia has a strong incentive to cultivate a reputation as the 'protector' of these minorities. Any intervention in one country will be observed by the other 13 countries, potentially leading to concessions. Of course, this strategy is viable only if the potential target countries take the Russian co-ethnic narrative seriously – even if they do not believe it reflects a genuine concern.

Why do target countries take this narrative seriously, even if they doubt Russia's genuine concern for Russian minorities abroad? Putin's Russia, being significantly stronger than most of its neighbors, has the capacity to inflict considerable harm through its proxies (large π and large C_T). As a result, even the unlikely prospect of Russian co-ethnic military support poses a substantial risk. Furthermore, as stated above, with a wide array of target countries, Russia has a strong incentive to cultivate a reputation as an 'intrinsic' type. Even if a target country perceives Russia as probably state-centric ('extrinsic' type), it will still pay attention to the Russian co-ethnic narrative and might opt to appease its Russian minority through political inclusion.[13]

Why is this region so prone to proxy wars involving Russian minorities? To answer this question, we have to look at Russian foreign policy before Putin took power. Russia's Compatriots policy can be traced back to the early 1990s,[14] but during this period, actions did not align with the stated policy objectives. Neighboring countries

---

[13]These two 'types' ascribed to the Russian state have counterparts in the two dominant Russian national identity ideologies. Indeed, the Russian language differentiates between *rossiikoe*, the citizens of the Russian state, and *russkoe*, those of Russian heritage regardless of their state. Laitin (1998) associates the first term with an *exclusive* national identity (state-centric or 'extrinsic'), and the second one with an *inclusive* ('intrinsic') national identity. Hence, the two 'types' of foreign sponsors are more than a mere theoretical simplification in the Russian case.

[14]The 1991 citizenship law granted Russian citizenship to co-ethnic groups stranded in the Near Abroad. Additionally, see Yeltsin's 1992 decree "On the Protection of the Rights and Interests of Russian Citizens outside the Russian Federation."

took notice. Shortly after gaining independence, Latvia and Estonia implemented discriminatory policies against their Russian minorities (Laitin, 1998). In response, Russia threatened intervention, notably to maintain Soviet troops stationed in these countries and to withhold economic aid or terminate preferential trade agreements, but never followed through on any of these threats. More importantly, Russia did not provide support to Russian Crimeans when "Crimea experienced the most intense movement for separatism and union with Russia." This lack of assistance led to the defeat of Crimean separatists in the 1998 election (Batta, 2021). In our model, these two events correspond to an extrinsically motivated state failing to intervene when challenged and hence revealing its true nature.

Thus, when Putin took power in 2000, he inherited a weak 'common prior belief' (low θ) regarding Russia's intrinsic motivation from the Yeltsin era. Nonetheless, the change in leadership plausibly gave him an opportunity to reset the country's reputation.[15] According to the model, a low θ ensures we are below the threshold for which target countries concede with certainty. Furthermore, large N, π, and C_T make it more likely that the sponsor state will intervene when there is no concession. This results in a situation especially prone to conflicts over reputation concerns. On the one hand, target countries are suspicious of Russia's true intention and are unlikely to concede without testing its resolve. On the other hand, because Russia has a large audience of target countries and the ability to intimidate them through its strength, it has a strong incentive to defend its reputation. This leads to an explosive situation and may explain the multiplicity of proxy wars involving local Russian populations in the 21st century.

3   Empirical strategy

3.1   Preliminary considerations

In order to find empirical support for reputation-building in international relations, through the threat of supporting proxy wars, we take the following precautions.

First, to limit the extent to which sponsors can manipulate narratives for lending credibility to their threats and justifying interventions, and to define clear sets of potential targets for each sponsor, we focus on their co-ethnicity with foreign rebel groups.

Sponsors also mobilize political or religious narratives. As far as the former is concerned, 10% of the interventions reported by San-Akca (2016) between 1945 and 2010 targeted groups displaying a socialist ideology (see also Bernauer, 2016). Besides,

---

[15]The question of whether reputations adhere to countries or leaders remains open (Jervis, 1982, Huth, 1997, and Dafoe et al., 2014). Renshon et al. (2018) provide evidence that country-specific reputations tend to be more prominent than leader-specific reputations, but they do not overshadow them completely.

Berman & Lake (2019) show that the United States has often relied on a pro-democracy narrative to justify its interventions through proxies. Moreover, Erstad (2018) shows that in its quest to establish itself as a regional power, the Shia regime in Iran supported Shia rebel groups in most of its neighbors,[16] suggesting that religion is another convincing narrative to support the threat of intervention. However, conflicts caused by ethnic divisions still represent a large sub-sample of all intrastate conflicts (Horowitz, 2000, Brubaker & Laitin, 1998, Fearon & Laitin, 2003, Esteban & Ray, 2008, Esteban et al., 2012, Rohner et al., 2013, Morelli & Rohner, 2015, Michalopoulos & Papaioannou, 2016, and Spolaore & Wacziarg, 2016), and support to co-ethnic groups a large proportion of sponsorship interventions: 33% of those reported by San-Akca (2016) over the 1945–2010 period.

Among all possible narratives, we argue that transnational co-ethnic ties are least likely to be manipulated by the sponsor state for reputation-building concerns or by rebel groups in the target states looking for funding. This does not mean that we are adopting an 'essentialist' or 'primordialist' view of ethnicity. Important works in sociology, political science, and economics, show that ethnicity is socially constructed (Bates, 1974, 1983, Hobsbawm & Ranger, 1983, Brass, 1991, Robinson, 2001, Chandra, 2006, Varshney, 2007, Eifert et al., 2010, Esteban et al., 2012, Vermeersch, 2012, Caselli & Coleman, 2013, Bisin et al., 2016, and Berman et al., 2023). Most of these works acknowledge that while ethnic partitions are largely exogenous and persistent, the relative salience of ethnicity is instrumentalized. In fact, this instrumentalization of ethnicity is an integral part of using co-ethnicity to build a sponsor's reputation (Sambanis et al., 2020).

Second, our empirical validation of Result 2 relies on the conservative subset of cases where concessions take the form of political representation for the ethnic group in the target state. The possibility of concessions unrelated to the co-ethnic group's political representation implies that the number of co-ethnic ties above which the sponsor state obtains an automatic concession is lower (*i.e.*, situation a persists for more stages in equilibrium). This limits the number of cases on which we can validate the argument (Tremblay-Auger, 2019).

Third, we do not know whether sponsor states are budget-constrained or not when it comes to financing rebellions. As explained in Section 2.4 above, this only makes our results harder to test empirically, and our approach more conservative.

3.2 Data

We rely on the Ethnic Power Relations (EPR) data set family to characterize the structure of ethnic ties between countries (Vogt et al., 2015). In particular, we use the EPR Core

---

[16]Although, it also supported the Alaouite and secular Baas regime in Syria and the Sunni Hezbollah.

2019 data set, which covers every country in the world between 1946 and 2017, to identify politically relevant ethnic groups in every potential target and sponsor state and characterize their level of political representation.[17] We spot groups that are present in both a target and a sponsor state thanks to the Transborder Ethnic Kin 2019 data set, which records all co-ethnic groups across countries. On the target state side, we rely on the ACD2EPR 2019 data set (Wucherpfennig et al., 2012 and Vogt et al., 2015, based on the UCDP/PRIO Armed Conflict data set by Gleditsch et al., 2002 and Pettersson & Eck, 2018) to relate existing rebel groups to ethnic groups. To make sure that the rebel group – ethnic group match is precise, we focus on rebel groups with a unique ethnic attachment, and leave groups with a mixed ethnic background aside.[18]

To characterize the sponsoring of a rebel group by another state, we rely on the Dangerous Companions: Cooperation between States and Non-state Armed Groups (NAGS) data set (San-Akca, 2016), which systematically covers cases of international support of rebel groups between 1925 and 2010.

Our unit of observation is the 'triad-year': an ethnic group, a target state and a potential sponsor, in a given year. An ethnic group is included as soon as it is politically relevant – as defined by the EPR Core classification – in the target state. We consider as potential sponsors for each target state the set of contiguous countries, as established in the Direct Contiguity data set (v3.2) of the Correlates of War project (Stinnett et al., 2002).[19] In Section 4.2.4, we consider an alternative definition of neighborhood based on macro-regions.

Contiguity serves as an objective criterion to keep the size of our data set manageable. Furthermore, we argue that co-ethnic links are more likely to be politically salient between neighbors. In fact, the "ethnic minority" question is less likely to be at the center of the relations between two distant countries, even if such an ethnic connection exists (Huth, 1999). In particular, the political integration of a foreign co-ethnic group is not prone to be of political interest; the home audience often cares less and the resulting alignment in policies is less likely to be useful. It is then arguably harder to develop a reputation for caring about the political inclusion of co-ethnics in distant polities and to use this reputation to obtain concessions. Moreover, general contentious issues between

---

[17]"An ethnic group is considered politically relevant if at least one political organization has claimed to represent its interests at the national level or if its members are subjected to state-led political discrimination." For definitions of ethnicity, representation, discrimination, and other important aspects of the construction of the data set, see Vogt et al. (2015) and the EPR Core Dataset Codebook, version 2019. The definitions of the political statuses that are used in the analysis are reported in Appendix A.

[18]For robustness, rebel groups with multiple ethnic background are considered along with single-ethnicity groups in Appendix C.c..

[19]We rely on the broadest definition of contiguity, which includes both land and water contiguity below 400 miles.

non-neighbors are less salient on average than between neighbors (Vasquez, 1995).[20]

The resulting data set is an unbalanced panel of 280,300 triad-year observations, reflecting 6,960 triads that gather 176 potential sponsors, 145 target states, and 819 ethnic groups over the 1946-2010 period. Only about a quarter of the triads are present for all 65 years under analysis due to variation in the recognition of politically relevant ethnic groups in the EPR Core data set.

Table 6 in Appendix A displays the descriptive statistics in our sample at the triad × year level, sponsor × year level, and group × year level.

About 4% of our triad × year observations are characterized by an ethnic tie, and 0.3% by an active support from the sponsor to a rebel group of the considered ethnicity in the target country. Among triads with an ethnic tie, the latter figure exceeds 2%: as expected, active support is thus much more common between co-ethnic actors.

Collapsing our sample at the sponsor × year level, we observe that nearly half of the yearly observations of potential sponsors are characterized by (at least) one ethnic tie with a foreign group, and 8% by an active support of a rebellion abroad (more than 14% if we focus on potential sponsors with at least one ethnic link with a group abroad). Moreover, among the 176 potential sponsor states of our sample, more than a quarter (47) have actively supported a foreign group at least once over the period (figure not reported in the table).

On the other hand, when we collapse our data at the group × year level, it appears that nearly 22% of the observations are characterized by at least one link with a potential sponsor in the neighborhood, and nearly 2% by a support received from abroad (more than 2.5% among groups with an ethnic tie). Notice also that, on average, groups that are co-ethnic with neighboring countries represent a larger share of the population in the target state than groups that are not. They are also more frequently in favorable political positions in the target country.

---

[20]For analogous arguments in favor of focusing on co-ethnic groups in contiguous countries, see Cederman et al. (2009, 2013); San-Akca (2016). Most conspicuously, using the contiguity criterion to define the set of potential sponsors eliminates instances of the United States supporting rebel groups in the Sudan, Rwanda, Angola, Nicaragua, El Salvador, Indonesia, Afghanistan, and many other countries; of the United Kingdom in various African, Middle-Eastern, and Asian countries, etc. Such instances often respond to other criteria than co-ethnicity. Nevertheless, we do lose some relevant links, such as Russia's with ethnic Russian minorities in Armenia, Kyrgyzstan, Moldova, Tajikistan, and Uzbekistan. With still twenty-two neighbors, and seven neighbors with co-ethnic ties, Russia remains one of the sponsors with the most potential targets in our data set.

3.3 Specifications

To test Result 1, we start by estimating the following specification:

$$S_{ijkt} = \alpha + \gamma L_{ijkt} + \delta N_{kt} + \eta L_{ijkt} \times N_{kt} + \phi z'_{ijt} + \chi x'_{ikt} + \lambda_{ijk} + \mu_t + \epsilon_{ijkt}, \quad (1)$$

where S_{ijkt} is a dummy variable that takes the value 1 when ethnic group i in country j receives active support for its rebellion from sponsor state k at time t. L_{ijkt} is the dummy capturing the existence of an ethnic tie involving group i between countries j and k: it takes the value 1 when ethnic group i, which is (by construction) politically relevant in country j, is in power in state k in year t. Specifically, i is considered to be in power in k if its status is *Monopoly*, *Dominant* or *Senior partner* (as opposed to groups that are either *Absent* or *Junior partner*, *Powerless*, *Self-excluded*, *Discriminated* against or *Irrelevant* in k).[21] N_{kt} is the number of neighbors of k where at least one ethnic group in power in k is politically relevant, at time t. In other words, it is the number of potential target states where country k's reputation could matter.[22]

We control for time-varying observable features z_{ijt} and x_{ikt} of ethnic group i in both target j and sponsor k states, that all come from the EPR Core data. In particular, we control for the status of the ethnic group in the target and in the sponsor countries, which is expected to affect the propensity of country k to support group i in country j. This amounts to introducing 6 dummies on the target side (*Monopoly*, *Dominant*, *Senior partner*, *Junior partner*, *Self-excluded* and *Powerless*, *Discriminated* being the omitted category) and 8 dummies on the sponsor side (*Monopoly*, *Dominant*, *Senior partner*, *Junior partner*, *Self-excluded*, *Powerless*, *Discriminated* and *Irrelevant*, *Absent* being the omitted category). In addition, z_{ijt} and x_{ikt} include quartile dummies for the size of i as a share of the population in countries j and k, to account for the respective weights that group i represent among the two populations.[23]

We also control for any time-invariant characteristic of the country pair jk, including all standard gravity controls (common language, history, etc.), and of ethnic group i within the country pair, with triad fixed effects [FE] λ_{ijk}. This allows us to only consider within-triad variations.[24] Last, we include year FE μ_t: they capture global time trends

---

[21]The definition of the various statuses as provided by the EPR data are reported in Appendix A.

[22]In a robustness test, we consider the total number of ethnic ties that k shares with its neighbors, rather than the number of countries with whom k shares at least one ethnic tie, as an alternative measure of the number of ethnic ties of the sponsor state. See Table 4 in Section 4.2.

[23]Controlling for the size of group i in country j partially captures a potential confounding effect driven by emigration from country k. Indeed, a larger diaspora from k could both induce a larger number of co-ethnic neighbors for k, and a larger group of co-ethnics – hence a larger propensity to intervene – in country j. To further check that this mechanism is unlikely to drive our results, we also verify that the benchmark results are stable when we control for the dummies for the size quartile of group i in country j along with their interaction with the link dummy L_{ijkt}. The results are available upon request.

[24]The triad FE thus also capture a potential confounding effect driven by cases of secession. Indeed,

16

and events likely to influence the probability of active support.

Coefficient γ in Equation 1 captures whether a group is more likely to receive support from a foreign state in which a co-ethnic group is in power. There is a unique source of within-triad variation in L_{ijkt}: when group i obtains or loses power in country k.[25] Thus, the set of dummies included in x_{ikt} that captures i's status in k at time t is collinear to L_{ijkt}, and we cannot estimate γ.

Coefficient δ captures whether a group is more likely to receive support from a foreign state that counts more co-ethnic neighbors. There are three reasons why N_{kt} may vary within a triad ijk: if the ruling coalition in country k changes with (a) group i or (b) another group i' obtaining or losing power, thus re-shuffling k's portfolio of co-ethnic neighbors; and (c) if the set of politically relevant groups in k's neighbors changes. Several mechanisms beyond reputation-building may generate a correlation between a change in the ruling coalition in country k (a and b) and k starting to support group i in country j, such as an electoral platform of interventionism. It is harder to imagine how an ethnic group becoming politically relevant in a third country (c) could affect the likelihood that sponsor k supports a rebellion of i in target j, if not because k seeks to build its reputation *vis-à-vis* third countries.

If it seeks to establish a reputation for resolve, a state with more co-ethnic neighbors is more likely to intervene in one of them, relative to a state with less co-ethnic neighbors. We test this prediction through the interacted term $L_{ijkt} \times N_{kt}$. A positive η would indicate that active support from a co-ethnic foreign country is more likely when the sponsor has more neighbors with co-ethnic groups, controlling for the direct effects of co-ethnicity and of a larger neighborhood. Our empirical approach to detect the existence of reputation-building is thus conceptually close to a difference-in-differences, in which we compare the probability of being sponsored by a state with many co-ethnic neighbors to the probability of being sponsored by a state with few co-ethnic neighbors, in target states that share a co-ethnic tie with the sponsor state relative to target states that do not.

We illustrate the three types of variations that are exploited in our empirical analysis with examples taken from the data where the number of co-ethnic neighbors changes for the potential sponsor state (namely, Afghanistan in 2006 and Thailand in 1980), in Appendix B.

In a second specification, we replace μ_t by target-sponsor-year FE ξ_{jkt}. This allows us to control for time-varying factors that may affect the relationship between target and

---

quite mechanically, each of the newly created countries after a secession is likely to count a large number of co-ethnic neighbors, while one could also expect a high propensity of intervention between countries that were formerly part of the same entity. In our setting, such a historical background is captured by the triad fixed effects.

[25]If group i becomes politically relevant in country j, the triad ijk appears in the data set, and if it ceases to be so, the triad drops out: neither would correspond to a within-triad variation of L_{ijkt}.

sponsor states, such as the signing of a new treaty or changes in the relative size of the two countries. This specification also accounts for any time-varying particulars of target and sponsor states, such as the time elapsed since an election or the occurrence of a natural disaster. We thus estimate:

$$S_{ijkt} = \alpha + \gamma L_{ijkt} + \delta N_{kt} + \eta L_{ijkt} \times N_{kt} + \phi z'_{ijt} + \chi x'_{ikt} + \lambda_{ijk} + \xi_{jkt} + \epsilon_{ijkt}. \qquad (2)$$

Notice that since N_{kt} varies with t at the level of the sponsor state k, we cannot estimate δ anymore when we include target-sponsor-year FE.

In our first two specifications, $L_{ijkt} \times N_{kt}$ may vary for several reasons: (a) it varies from 0 to N_{kt} if group i obtains power in country k at time t, and from N_{kt-1} to 0 if it loses power; (b) it varies from N_{kt-1} to N_{kt} if a group i' joins or leaves i in the coalition in power in k, thus leaving k with more or fewer co-ethnic neighbors; and (c) it varies from N_{kt-1} to N_{kt} if a change in the ethnic composition of a neighbor of k means that k gains or loses co-ethnic neighbors.

We estimate a third specification that neutralizes the first source of variation mentioned above. To do so, we replace our triad fixed effects by fixed effects for triad $\times$ power-in-k dummies, where power-in-k dummies indicate whether group i is part of the coalition in power in k, or not. Consistent with our definition of L_{ijkt}, a group is considered in power in k when its political status is *Monopoly*, *Dominant*, or *Senior partner*. We then exploit solely variations within triads in which i is either in or out of power in country k. The corresponding specification can be written as follows:

$$S_{ijkt} = \alpha + \gamma L_{ijkt} + \delta N_{kt} + \eta L_{ijkt} \times N_{kt} + \phi z'_{ijt} + \chi x'_{ikt} + \Lambda_{ijk} + \xi_{jkt} + \epsilon_{ijkt}, \qquad (3)$$

where Λ_{ijk} are power-in-k-specific triad fixed effects. With these, we use only the two latter sources of variation of $L_{ijkt} \times N_{kt}$ – specifically (b) another group joining i in power in k and (c) changes in the ethnic composition of k's neighbors – to estimate η. Ideally, we would focus solely on the last source of variation, but due to the similar structural changes that (b) and (c) imply in the data, we cannot eliminate (b) as we did with (a). However, we address it with additional controls in Section 4.2.1.

Finally, Result 1 suggests a nonmonotonic effect of $L_{ijkt} \times N_{kt}$ on the likelihood of an active support of k to i. We should find that $\eta > 0$ when N_{kt} is small enough, but when N_{kt} is larger, the theory suggests that country j would concede to country k rather than see k support a rebellion by group i with certainty, *i.e.* that $\mathbb{P}(S_{ijkt} = 1) = 0$. To examine such non-monotonicity in the data, we employ a semi-parametric approach, replacing N_{kt} in Equation 3 with dummy variables representing possible ranges of values for N_{kt}. In our sample of analysis, N_{kt} ranges between 0 and 9. We introduce four dummies,

$d(N=2,3)_{kt}$, $d(N=4,5)_{kt}$, $d(N=6,7)_{kt}$ and $d(N=8,9)_{kt}$, which take the value 1 when N_{kt} is equal to 2 or 3; 4 or 5; 6 or 7, and 8 or 9, respectively. The reference category is $N_{kt}=1$, while the category $N_{kt}=0$ is omitted due to perfect collinearity with L_{ijkt}. The resulting specification is as follows:

$$Y_{ijkt} = \alpha + \eta_1 L_{ijkt} \times d(N=2,3)_{kt} + \eta_2 L_{ijkt} \times d(N=4,5)_{kt}$$
$$+ \eta_3 L_{ijkt} \times d(N=6,7)_{kt} + \eta_4 L_{ijkt} \times d(N=8,9)_{kt}$$
$$+ \phi z'_{ijt} + \chi x'_{ikt} + \Lambda_{ijk} + \xi_{jkt} + \epsilon_{ijkt}. \quad (4)$$

To test Result 2, we estimate the same four specifications, using an indicator of the political inclusion of group i in country j at time t as dependent variable.[26] Specifically, I_{ijt} is a qualitative, ordinal variable built from the EPR Core data, equal to 0 if group i is *Discriminated*, *Powerless* or *Self-excluded* in country j; to 1 if it is a *Junior* or a *Senior partner* in j; and to 2 if it is *Dominant* or in a *Monopoly*. When using I_{ijt} as dependent variable, we only include the quartile dummies for the size of i as a share of the population in country in j in the vector of controls z_{ijt}, while x_{ikt} remains unchanged (*i.e.*, it includes quartile dummies for the size of i as a share of the population in country k and 8 dummies for the status of group i in k). Result 2 corresponds to $\eta > 0$, and the theory does not suggest any non-monotonicity there.

We estimate the four specifications using OLS, both for active support and political inclusion. Given the rich set of fixed effects included, this is our preferred estimator even though the dependent variables are qualitative.[27] This approach allows us to work with the complete sample and avoid potential selection bias arising from restricting the sample to switching observations. In the robustness section, we show that the OLS estimations yield reasonable predicted values.

Lastly, we cluster standard errors at both the sponsor and target levels. Clustering at the sponsor level allows its reputation to be correlated non-parametrically across time and across possible target states among its neighbors. Clustering at the target level allows its sensitivity to reputation to be correlated non-parametrically across time and across various ethnic groups through which a threat may emerge.

---

[26]Note that political inclusion of i is a feature of a country, not of a target-sponsor dyad. Using the triad-level specifications is however useful for comparability with the results on active support, and because it allows to isolate the reputation-building mechanism more convincingly. In the robustness section 4.2, we test an alternative specification on a data set collapsed at the ethnic group × target × year level.

[27]Timoneda (2021) shows that FE-LPM produces more accurate estimates than logistic regression with group intercepts or conditional logit when the dependent variable is binary and accounts for less than 25% of ones.

4 Empirical results

4.1 Benchmark

Columns (1)-(3) of Table 1 present the results of the benchmark specifications 1 to 3. In line with our theoretical Result 1, $\hat{\eta}$ is positive and significant. This indicates that if group i, relevant in country j, is a member of the ruling coalition in k, active support from country k to group i in country j is significantly more likely if k has more co-ethnic neighbors. This result remains robust when introducing target-sponsor-year fixed effects (Column (2)) and triad × power-in-k fixed effects (Column (3)). The effect is substantial: in Column (3), an additional co-ethnic neighbor is estimated to be associated with a nearly 1pp increase in the probability of sponsorship.

Columns (4)-(6) report the results of the benchmark estimations when I_{ijt} is the dependent variable. They are in line with our theoretical Result 2, as the estimated coefficient associated with $L_{ijkt} \times N_{kt}$ remains consistently positive across columns. This indicates that the political representation of group i in country j is significantly more important, on average, when co-ethnic neighboring countries k have more neighbors with shared ethnic group(s).

Figure 1 graphically presents the results of the semi-parametric specification 4, plotting the point estimates for the four coefficients of interest ($\hat{\eta}_1$, $\hat{\eta}_2$, $\hat{\eta}_3$ and $\hat{\eta}_4$) along with their 95% confidence intervals. While the overlapping confidence intervals preclude us from drawing strong conclusions on non-monotonicity, the comparison between the left-hand side and the right-hand side panels – where we use inclusion I_{jkt} as the dependent variable – suggests a different effect across dependent variables. This finding is consistent with the theoretical prediction that political representation should increase in the number of neighbors in which a potential foreign sponsor can sponsor rebellions, while the probability of active support should increase with this number until it reaches a sufficiently high threshold level.

4.2 Robustness

4.2.1 Changes in the ruling coalition in k

As discussed above, the source of variation of N_{kt} which impact can most plausibly be attributed to reputation-building is the change in the set of politically relevant groups in k's neighbours – by opposition to changes in the ruling coalition in k. In Equation 3 and the corresponding Columns (3) and (6) of Table 1, the introduction of triad × power-in-k dummies allows to neutralize the change in group i's political role as a source of variation of $L_{ijkt} \times N_{jkt}$. It can still be the case, however, that $L_{ijkt} \times N_{kt}$ changes not because

Table 1: Benchmark results

	(1)	(2)	(3)	(4)	(5)	(6)
		Active support			*Political inclusion*	
Nb of countries with a link	-0.0007			-0.0050		
	(0.001)			(0.003)		
	[0.175]			[0.056]		
Link × Nb of countries with a link	0.0056	0.0079	0.0096	0.0374	0.0610	0.0735
	(0.003)	(0.004)	(0.006)	(0.020)	(0.030)	(0.038)
	[0.081]	[0.036]	[0.084]	[0.055]	[0.040]	[0.056]
Controls in Target$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes
Controls in Sponsor$_{ikt}$	Yes	Yes	Yes	Yes	Yes	Yes
Triad$_{ijk}$ FE	Yes	Yes		Yes	Yes	
Year$_t$ FE	Yes			Yes		
Target$_j$*Sponsor$_k$*Year$_t$ FE		Yes	Yes		Yes	Yes
Triad$_{ijk}$ × Power-in-k FE			Yes			Yes
R^2	0.4706	0.5993	0.6037	0.9237	0.9341	0.9349
Adjusted R^2	0.4570	0.5194	0.5244	0.9218	0.9209	0.9219

Notes: OLS estimations at the triad level. 280,300 observations (corresponding to 145 target states, 176 sponsor states, 819 groups). Controls target state include four quartile dummies for the group size among the total population in all six columns, and six political status dummies in Columns (1) to (3) (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless* and *Self-excluded*; with *Discriminated* being the group of reference). Controls sponsor state include four quartile dummies for the group size among the total population and eight political status dummies in all six columns (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless*, *Self-excluded*, *Discriminated* and *Irrelevant*, with *Absent* being the group of reference). Robust standard errors clustered at the target-country and sponsor-country levels in parentheses. P-values in squared parentheses.

Figure 1: Documenting possible non-monotonicity

Dependent variable: Active support   Dependent variable: Political inclusion
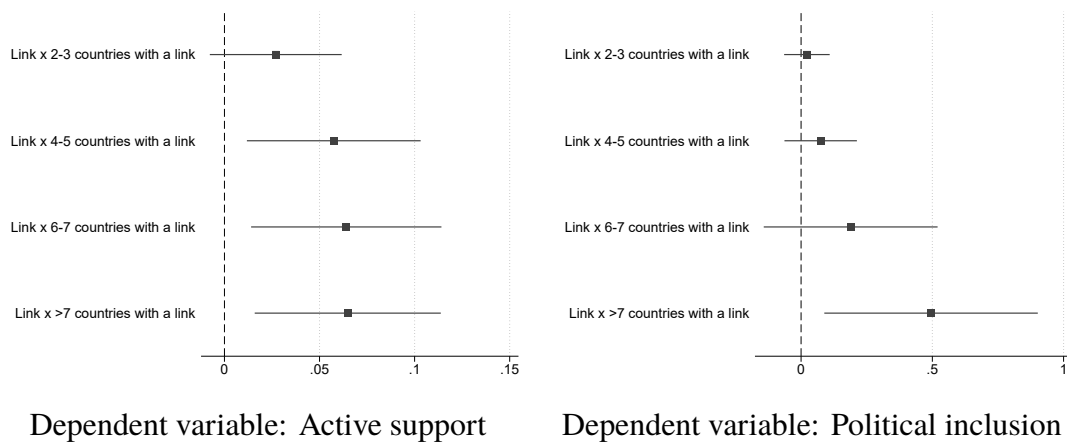
of a change in the ethnic structure of k's neighboring countries, but due to group(s) i' joining or leaving i in the ruling coalition in k, which would induce a reshuffle in k's portfolio of co-ethnic neighbors.

To account for this possibility empirically, in Table 2 we reproduce the specifications of Table 1 additionally controlling for a variable measuring the number of groups in power in k (G_{kt}) and its interaction with L_{ijkt}. Since the number of groups in power can remain constant while the composition of the coalition changes, we also control for a dummy that flags cases of power switches in k with no variation in the number of ruling groups (PS_{kt}), and its interaction with L_{ijkt}. Notice that, as N_{kt}, G_{kt} and PS_{kt} vary at the sponsor-year level and are thus perfectly collinear to our target-sponsor-year FE, their direct coefficients cannot be estimated in specifications 2 and 3.

The results are reassuring regarding the potential impact of changing coalitions in k on the estimated coefficient for $L_{ijkt} \times N_{kt}$. They show that neither a change in the number of groups in power in k nor a switch in the composition of a constant-size coalition is associated with significantly different probabilities of support, irrespective of whether group i is in power in country k or not. Furthermore, our coefficient of interest remains quite stable in both magnitude and statistical significance when we incorporate these additional controls.

4.2.2 Data geometry

Our benchmark estimations exploit a triad-level database. We develop three tests to verify that this specific format of the data is not driving our main results.

First, unlike S_{ijkt}, I_{ijt} is group- and target country-specific but does not depend on k. It is useful to run the same specifications on our two outcomes of interest for comparability, and it is consistent with the model to build our explanatory variable of interest (*i.e.*, $L_{ijkt} \times N_{kt}$) at the level of sponsor states k, as in Table 1 and Figure 1. However, using the triad as the unit of observation to estimate the determinants of political inclusion in country i could introduce bias if, for instance, countries with more neighbors – thus, observed in more triads – follow specific political trends.

To validate our results on political inclusion in a more natural, group-level setting, we build an alternative collapsed data set, using ethnic group i in country j – rather than triads between ethnic group i, country j, and country k – as the unit of observation. We then estimate the following model:

$$I_{ijt} = \alpha + \kappa K_{ijt} + \nu N K_{ijt} + \phi z'_{ijt} + \lambda_{ij} + \mu_t + \epsilon_{ijt}, \tag{5}$$

where K_{ijt} is the number of possible sponsor states where group i is in power, and NK_{ijt} is the sum of the number of countries with whom each of these possible sponsor states

Table 2: Changes in the ruling coalition

	(1)	(2)	(3)	(4)	(5)	(6)
	Active support			*Political inclusion*		
Nb of countries with a link	-0.0007			-0.0049		
	(0.001)			(0.003)		
	[0.226]			[0.064]		
Link × Nb of countries with a link	0.0058	0.0086	0.0098	0.0372	0.0596	0.0724
	(0.003)	(0.004)	(0.005)	(0.020)	(0.030)	(0.039)
	[0.088]	[0.039]	[0.066]	[0.063]	[0.051]	[0.065]
Nb of groups in power	0.0006			-0.0014		
	(0.001)			(0.004)		
	[0.634]			[0.746]		
Link × Nb of groups in power	-0.0033	-0.0106	-0.0046	0.0052	0.0207	0.0183
	(0.014)	(0.020)	(0.040)	(0.021)	(0.032)	(0.043)
	[0.817]	[0.603]	[0.908]	[0.807]	[0.512]	[0.669]
Power switch	-0.0006			-0.0056		
	(0.001)			(0.005)		
	[0.558]			[0.248]		
Link × Power switch	-0.0161	-0.0215	-0.0202	0.0383	0.0523	0.0120
	(0.011)	(0.014)	(0.016)	(0.057)	(0.066)	(0.042)
	[0.144]	[0.114]	[0.201]	[0.504]	[0.432]	[0.775]
Controls in Target$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes
Controls in Sponsor$_{ikt}$	Yes	Yes	Yes	Yes	Yes	Yes
Triad$_{ijk}$ FE	Yes	Yes		Yes	Yes	
Year$_t$ FE	Yes			Yes		
Target$_j$*Sponsor$_k$*Year$_t$ FE		Yes	Yes		Yes	Yes
Triad$_{ijk}$ × Power-in-k FE			Yes			Yes
R^2	0.4706	0.5994	0.6037	0.9237	0.9341	0.9349
Adjusted R^2	0.4570	0.5194	0.5244	0.9218	0.9209	0.9219

Notes: OLS estimations at the triad level. 280,300 observations (corresponding to 145 target states, 176 sponsor states, 819 groups). Controls target state include four quartile dummies for the group size among the total population in all six columns, and six political status dummies in Columns (1) to (3) (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless* and *Self-excluded*; with *Discriminated* being the group of reference). Controls sponsor state include four quartile dummies for the group size among the total population and eight political status dummies in all six columns (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless*, *Self-excluded*, *Discriminated* and *Irrelevant*, with *Absent* being the group of reference). Robust standard errors clustered at the target-country and sponsor-country levels in parentheses. P-values in squared parentheses.

shares an ethnic link. We control for group and year fixed effects λ_{ij} and μ_t, which account for time-invariant characteristics of the group that may correlate with their level of political representation and for global time trends, respectively. We also control for the vector z_{ijt}, which includes quartile dummies for the size of i in the total population of country j and quartile dummies for the average size of group i in the total population in potential sponsor states. The results of this estimation are shown in Column (1) of Table 3. In Column (2), we replace μ_t by country-year fixed effects, that account for country-level time trends.

The results of the group-level analysis show that ethnic groups with more potential sponsors tend to have lower political representation on average. However, in line with our benchmark triad-level findings, the estimated coefficient for NK_{ijt} is significantly positive. This means that when accounting for the trans-border nature of an ethnic group, its political representation increases if potential foreign sponsors share more ethnic links with their neighbors. This suggests that the more opportunities sponsors have to support rebellions in neighboring countries, the greater the political representation of the trans-border ethnic group in its home country.

Table 3: Political inclusion, group-level estimations

Dependent variable: Political inclusion	(1)	(2)
Nb of possible sponsors with a link	-0.0630	-0.1141
	(0.048)	(0.058)
	[0.194]	[0.050]
Number of links of possible sponsors	0.0197	0.0351
	(0.011)	(0.015)
	[0.063]	[0.018]
Group$_{ij}$ FE	Yes	Yes
Year$_t$ FE	Yes	
Country$_j$*Year$_t$ FE		Yes
Controls$_{ijt}$	Yes	Yes
R^2	0.8988	0.9132
Adjusted R^2	0.8961	0.8889
Observations	34,844	34,843

Notes: OLS estimations at the ethnic group – target state level. Controls include four quartile dummies for the group size among the total population and four quartile dummies for the average group size among the total population in potential sponsor states. Robust standard errors clustered at the target-country level in parentheses. P-values in squared parentheses.

24

Another potential issue arising from the triad-level construction of our data set is that sponsor states with more neighbors are mechanically part of more triads. To address this, we weight the triads so that each sponsor state is assigned a weight of 1, while imposing that all triads involving the same sponsor state (resp. all years involving the same triad) have the same weight for the sponsor. The results of the weighted estimations, displayed in Table 7 in Appendix C.a., are close to our benchmark findings.

To provide further evidence that the triad dimension of our benchmark data set does not drive our results, we collapse the data at the country-pair level. It should be noted that this exercise (i) only addresses the fact that a mechanical inflation in the number of observations happens for sponsor states whose target neighbors have more politically relevant groups, but not the fact that it also happens for those with more target neighbors; and (ii) is only performed for the active support dependent variable. The results of this test, presented in Table 8 in Appendix C.b., are in line with those of Table 1.

4.2.3 Alternative measures

First, we explore the robustness of our benchmark results to the use of two alternative measures of the number of ethnic ties of the sponsor state.

Our benchmark estimates (Table 1) rely on the number of countries with whom k shares at least one ethnic tie – a variable which, schematically, reflects in how many places k's reputation may matter. In Panel A of Table 4, we count the total number of ethnic ties that k shares with its neighbors instead. We thus introduce another margin of intensity, which authorizes the reputation-building at stake in k's active support to group i in country j to be also more important if k shares ethnic links with other groups in country j, and/or if k shares more numerous ethnic links with the same set of neighboring countries. In Panel B of Table 4, we count the number of countries with whom k shares an i-type ethnic tie – the idea being here that the reputation of k may be (partly) group-specific, *i.e.* a sponsor state may have the reputation to defend specific groups. In both cases, the results are consistent with our benchmark findings, especially once our complete set of fixed effects is introduced. Nonetheless, the coefficients of interest are larger in Panel B, which suggests that the reputation of k is at least partly group-specific.

Secondly, all the results presented above focus on rebel groups for which the ACD2EPR data identify one single ethnic background. This approach minimizes the risk of misclassification when relating rebel group-level violence data to ethnic group-level political data, but it comes at the cost of disregarding part of the available information. As an additional robustness test, we run the same estimations as in Table 1 accounting for all the ethnicity data provided by the ACD2EPR. If a sponsored rebel group is related

25

Table 4: Alternative measures of the number of links

	(1)	(2)	(3)	(4)	(5)	(6)
	Active support			*Political inclusion*		
Panel A						
Nb of links	-0.0006			-0.0050		
	(0.000)			(0.002)		
	[0.234]			[0.035]		
Link × Nb of links	0.0042	0.0057	0.0076	0.0291	0.0499	0.0577
	(0.003)	(0.003)	(0.005)	(0.016)	(0.024)	(0.030)
	[0.106]	[0.056]	[0.098]	[0.073]	[0.037]	[0.054]
R^2	0.4706	0.5993	0.6037	0.9237	0.9341	0.9349
Adjusted R^2	0.4569	0.5194	0.5244	0.9218	0.9210	0.9219
Panel B						
Nb of countries with an i-type link	-0.0844	-0.1080	-0.0038	-0.1206	-0.1300	-0.1853
	(0.061)	(0.080)	(0.005)	(0.129)	(0.109)	(0.129)
	[0.170]	[0.177]	[0.478]	[0.349]	[0.235]	[0.151]
Link × Nb of countries with an i-type link	0.0899	0.1172	0.0144	0.1567	0.1929	0.2699
	(0.061)	(0.081)	(0.008)	(0.125)	(0.109)	(0.131)
	[0.142]	[0.148]	[0.074]	[0.210]	[0.077]	[0.039]
R^2	0.4711	0.6000	0.6037	0.9237	0.9341	0.9349
Adjusted R^2	0.4575	0.5202	0.5244	0.9217	0.9209	0.9219
Controls in Target$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes
Controls in Sponsor$_{ikt}$	Yes	Yes	Yes	Yes	Yes	Yes
Triad$_{ijk}$ FE	Yes	Yes		Yes	Yes	
Year$_t$ FE	Yes			Yes		
Target$_j$*Sponsor$_k$*Year$_t$ FE		Yes	Yes		Yes	Yes
Triad$_{ijk}$ × Power-in-k FE			Yes			Yes

Notes: OLS estimations at the triad level. 280,300 observations (corresponding to 145 target states, 176 sponsor states, 819 groups). Controls target state include four quartile dummies for the group size among the total population in all six columns, and six political status dummies in Columns (1) to (3) (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless* and *Self-excluded*; with *Discriminated* being the group of reference). Controls sponsor state include four quartile dummies for the group size among the total population and eight political status dummies in all six columns (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless*, *Self-excluded*, *Discriminated* and *Irrelevant*, with *Absent* being the group of reference). Robust standard errors clustered at the target-country and sponsor-country levels in parentheses. P-values in squared parentheses.

to more than one ethnicity, we assume that all the related ethnic groups benefit from the considered external support. The results, shown in Table 9 in Appendix C.c., are consistent with our benchmark findings.

4.2.4 Defining the set of neighbors

As discussed in Section 3.2, our benchmark results are obtained over a sample that covers most countries of the world as potential target states, and all targets' contiguous neighbors as potential sponsors. While such a methodological choice is undoubtedly restrictive, we argue that it is consistent with our focus on co-ethnicity as narrative on which the reputation-building mechanism can be activated. Still, one could be concerned that this approach affects our results.

Therefore, in Table 5, we consider a more inclusive definition of neighborhood which is based on macro-regions. More specifically, each state is assumed to be a potential target for all the other countries in the same continent. The results prove very stable, as compared to those in Table 1, regardless of which dependent variable (*Active support* or *Political inclusion*) is considered.

4.2.5 Falsification test

We then re-estimate our benchmark specifications using a dependent variable that should not be reactive to reputation-building strategies: the passive support of sponsor k for a rebellion of group i in country j. The dummy variable for passive support comes from San-Akca (2016), and takes value 1 in cases where a rebel group uses a foreign state for its activity without the latter's consent. For instance, groups from country j recruiting combatants in state k, or hiding in state k are considered by San-Akca (2016) as benefiting from passive support. It is important to note that this placebo test is quite demanding, as (i) active and passive supports can go hand-in-hand (the coefficient of correlation between the two variables is equal to 34% in our sample), and (ii) if one assumes that states are capable of impeding it, passive support could then activate a low-intensity reputation-building mechanism (by building sponsor k's reputation, not as a rebellion-supporter, but as a rebellion-facilitator). Nevertheless, the results, shown in Table 10 in Appendix C.d., yield a non-significant coefficient for the interaction of interest, and a point estimate two to five times smaller compared to Columns (1)-(3) of Table 1.

Table 5: Alternative set of neighbors

	(1)	(2)	(3)	(4)	(5)	(6)
	Active support			Political inclusion		
Nb of countries with a link	0.0000			-0.0030		
	(0.000)			(0.001)		
	[0.762]			[0.000]		
Link × Nb of countries with a link	0.0030	0.0044	0.0077	0.0225	0.0336	0.0459
	(0.002)	(0.002)	(0.004)	(0.011)	(0.016)	(0.024)
	[0.134]	[0.057]	[0.032]	[0.034]	[0.034]	[0.054]
Controls in Target$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes
Controls in Sponsor$_{ikt}$	Yes	Yes	Yes	Yes	Yes	Yes
Triad$_{ijk}$ FE	Yes	Yes		Yes	Yes	
Year$_t$ FE	Yes			Yes		
Target$_j$*Sponsor$_k$*Year$_t$ FE		Yes	Yes		Yes	Yes
Power in sponsor$_{ik}$ FE			Yes			Yes
R^2	0.4476	0.5829	0.5878	0.8867	0.9034	0.9036
Adjusted R^2	0.4315	0.4646	0.4708	0.8833	0.8760	0.8763

Notes: OLS estimations at the triad level. 1,273,132 observations (corresponding to 145 target states, 184 sponsor states, 819 groups). Controls target state include four quartile dummies for the group size among the total population in all six columns, and six political status dummies in Columns (1) to (3) (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless* and *Self-excluded*; with *Discriminated* being the group of reference). Controls sponsor state include four quartile dummies for the group size among the total population and eight political status dummies in all six columns (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless*, *Self-excluded*, *Discriminated* and *Irrelevant*, with *Absent* being the group of reference). Robust standard errors clustered at the target-country and sponsor-country levels in parentheses. P-values in squared parentheses.

4.2.6 Estimation procedure

As discussed in Section 3, all the results displayed above are obtained with OLS. This can rise concerns related to the qualitative nature of our dependent variables. As a first check, we compute predicted values from Table 1. As shown in Table 11 in Appendix C.e., few predicted values fall far away from the [0,1] interval for active support, and from the [0,2] interval for political inclusion. Then, we re-estimate the benchmark specifications over the sub-sample of observations that yield predicted values within the [0,1] interval for active support, and the [0,2] interval for political inclusion. The results are reassuringly consistent with our benchmark findings (see Table 12 in Appendix C.e.).

5 Conclusion

The theoretical case for reputation-building is well-established in international relations. However, for the notion to make sense at all, the true motives of the aggressor must be hidden. In fact, they cannot even be credibly communicated – or an extrinsically-motivated aggressor could not hope to misrepresent itself. Of course, hidden motives pose a unique empirical challenge. To tackle this difficulty, we have examined two non-trivial properties of the mechanism. While they cannot reveal a particular state's true motives, we can test empirically whether states behave in a way that is statistically consistent with reputation-building.

In particular, we have considered whether states may establish their reputation by threatening to sponsor rebellions of co-ethnic groups abroad, in states that refuse to make concessions. Based on exceptionally rich data at the level of the triads (ethnic group × target country × potential sponsor country), and using a demanding empirical strategy notably relying on variations in the ethnic structure of third countries, our results are consistent with states establishing a reputation for resolve by supporting foreign rebellions. Indeed, we find that a state is more likely to support a co-ethnic rebellion in a neighboring country if it has more co-ethnic neighbours, and that groups whom potential sponsor states count more co-ethnic neighbours also benefit from more political representation, on average.

The Russian case study highlights some important implications of our theory for the international community and policy-makers. Although one may be tempted to work to discredit a narrative such as the "Compatriots" policy, this may make matters worse by incentivizing the sponsor state to defend its reputation by supporting a rebellion. Instead, promoting the political inclusion of the co-ethnics and strengthening the target country militarily is more likely to deter conflict and provide the latter with a better bargaining position on the international stage.

While our results provide a compelling argument in support of reputation-building in international relations, more work needs to be done to understand the specific circumstances, timing, and mechanisms through which reputation-building unfolds. A key avenue for achieving this lies in incorporating insights from cultural and psychological approaches, alongside conducting further rigorous empirical tests.

References

Acemoglu, D. (2003). Why not a political Coase theorem? Social conflict, commitment, and politics. *Journal of Comparative Economics*, 31(4), 620–52.

Ahmadian, H. & Mohseni, P. (2019). Iran's Syria strategy: The evolution of deterrence. *International Affairs*, 95(2), 341–364.

Allison, R. (2013). *Russia, the West, and Military Intervention*. Oxford University Press.

Ansorg, N. (2011). How does militant violence diffuse in regions? Regional conflict systems in international relations and peace and conflict studies. *International Journal of Conflict and Violence*, 5(1), 173–87.

Baird, I. G. (2010). The construction of 'indigenous peoples' in Cambodia. In L. Yew (Ed.), *Alterities in Asia: Reflections on identity and regionalism* (pp. 155–176). Routledge.

Bak, D., Chávez, K., & Rider, T. (2020). Domestic political consequences of international rivalry. *Journal of Conflict Resolution*, 64(4), 703–28.

Balch-Lindsay, D., Enterline, A. J., & Joyce, K. A. (2008). Third-party interventions and the civil war process. *Journal of Peace Research*, 45(3), 345–63.

Bale, J. M. (2012). Terrorists as state "proxies": Separating fact from fiction. In M. A. Innes (Ed.), *Making Sense of Proxy wars: States, Surrogates & the Use of Force* (pp. 1–30). Potomac Books.

Bapat, N. A. (2007). The internationalization of terrorist campaigns. *Conflict Management and Peace Science*, 24(4), 265–80.

Bapat, N. A. (2012). Understanding state sponsorship of militant groups. *British Journal of Political Science*, 42(1), 1–29.

Barnhart, J. (2021). The consequences of defeat: The quest for status and morale in the aftermath of war. *Journal of Conflict Resolution*, 65(1), 195–222.

Başer, E. T. (2021). Reputations in international conflict and cooperation. `https://ekrembaser.com/pdf/reputation.pdf`. Retrieved on May 20th, 2021.

Bates, R. (1974). Ethnic competition and modernization in contemporary Africa. *Comparative Political Studies*, 6(4), 457–84.

Bates, R. (1983). Modernization, ethnic competition, and the rationality of politics in contemporary Africa. In D. Rothchild & V. A. Olorunsola (Eds.), *States Versus Ethnic Claims: African Policy Dilemmas* (pp. 153–71). Westview Press.

Batta, A. (2021). *The Russian Minorities in the Former Soviet Republics: Secession, Integration, and Homeland*. Routledge.

Bendix, A. (2022). Constructive role ambiguity and how Russia couldn't 'get away' with its 2022 Ukrainian invasion. *Central European Journal of International and Security Studies*, 16(3), 108–130.

Berman, E. & Lake, D. A. (2019). *Proxy Wars: Suppressing Violence Through Local Agents*. Cornell University Press.

Berman, N., Couttenier, M., & Girard, V. (2023). Natural resources and the salience of ethnic identities. *The Economic Journal*. Forthcoming.

Bernauer, E. (2016). *Identities in civil conflict*. Springer.

Bisin, A., Patacchini, E., Verdier, T., & Zenou, Y. (2016). Bend it like Beckham: Ethnic identity and integration. *European Economic Review*, 90, 146–64.

Braithwaite, A. (2010). Resisting infection: How state capacity conditions conflict contagion. *Journal of Peace Research*, 47(3), 311–9.

Brass, P. R. (1991). *Ethnicity and Nationalism: Theory and Comparison*. Sage publications.

Brubaker, R. & Laitin, D. D. (1998). Ethnic and nationalist violence. *Annual Review of Sociology*, 24, 423–52.

Buhaug, H. & Gleditsch, K. S. (2008). Contagion or Confusion? Why Conflicts Cluster in Space. *International Studies Quarterly*, 52(2), 215–33.

Byman, D. & Kreps, S. E. (2010). Agents of destruction? Applying principal-agent analysis to state-sponsored terrorism. *International Studies Perspectives*, 11(1), 1–18.

Caselli, F. & Coleman, W. J. (2013). On the theory of ethnic conflict. *Journal of the European Economic Association*, 11(s1), 161–92.

Cederman, L.-E., Girardin, L., & Gleditsch, K. S. (2009). Ethnonationalist triads: Assessing the influence of kin groups on civil wars. *World Politics*, 61(3), 403–37.

Cederman, L.-E., Gleditsch, K. S., Salehyan, I., & Wucherpfennig, J. (2013). Transborder ethnic kin and civil war. *International Organization*, 67(2), 389–410.

Chandra, K. (2006). What is ethnic identity and does it matter? *Annual Review of Political Science*, 9, 397–424.

Clare, J. & Danilovic, V. (2010). Multiple audiences and reputation building in international conflicts. *Journal of Conflict Resolution*, 54(6), 860–82.

Cormac, R. & Aldrich, R. J. (2018). Grey is the new black: Covert action and implausible deniability. *International Affairs*, 94(3), 477–94.

Dafoe, A., Renshon, J., & Huth, P. (2014). Reputation and status as motives for war. *Annual Review of Political Science*, 17(1), 371–93.

Eifert, B., Miguel, E., & Posner, D. N. (2010). Political competition and ethnic identification in Africa. *American Journal of Political Science*, 54(2), 494–510.

Ellington, T. C. (2003). Won't get fooled again: The paranoid style in the national security state. *Government and Opposition*, 38(4), 436–55.

Erstad, H. U. (2018). Transnational Mobilisation of Armed Resistance: How Iran Has Built a Network of Iraqi Shi'a Militias Since 1979. Master's thesis, University of Olso.

Esteban, J., Mayoral, L., & Ray, D. (2012). Ethnicity and conflict: Theory and facts. *Science*, 336(6083), 858–65.

Esteban, J. & Ray, D. (2008). On the salience of ethnic conflict. *American Economic Review*, 98(5), 2185–202.

Farrow, R. (2018). *Shadow Armies: Political Representation and Strategic Reality in America's Proxy Wars*. PhD thesis, University of Oxford.

Fearon, J. D. (1995). Rationalist explanations for war. *International Organization*, 49(3), 379–414.

Fearon, J. D. & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1), 75–90.

Fudenberg, D. & Kreps, D. M. (1987). Reputation in the simultaneous play of multiple opponents. *The Review of Economic Studies*, 54(4), 541–68.

Gates, S., Graham, B. A. T., Lupu, Y., Strand, H., & Strøm, K. W. (2016). Power sharing, protection, and peace. *The Journal of Politics*, 78(2), 512–26.

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed conflict 1946–2001: A new dataset. *Journal of Peace Research*, 39(5), 615–37.

Grigas, A. (2016). *Beyond Crimea: the New Russian Empire*. Yale University Press.

Hobsbawm, E. & Ranger, T. (1983). *The Invention of Tradition*. Cambridge University Press.

Horowitz, D. L. (2000). *Ethnic Groups in Conflict, Updated Edition with a New Preface*. University of California Press.

Huth, P. K. (1997). Reputations and deterrence: A theoretical and empirical assessment. *Security Studies*, 7(1), 72–99.

Huth, P. K. (1999). Deterrence and international conflict: Empirical findings and theoretical debates. *Annual Review of Political Science*, 2(1), 25–48.

Ibrahimi, N. (2017). *The Hazaras and the Afghan state: Rebellion, exclusion and the struggle for recognition*. Oxford University Press.

Jervis, R. (1982). Deterrence and perception. *International security*, 7(3), 3–30.

Kathman, J. D. (2010). Civil war contagion and neighboring interventions. *International Studies Quarterly*, 54(4), 989–1012.

Kathman, J. D. (2011). Civil war diffusion and regional motivations for intervention. *Journal of Conflict Resolution*, 55(6), 847–76.

Kausch, K. (2017). Proxy agents: State and non-state alliances in the Middle-East. *The International Spectator*, 52(3), 36–47.

Kibbe, J. D. (2011). Conducting shadow wars. *Journal of National Security Law and Policy*, 5(2), 373–92.

König, M. D., Rohner, D., Thoenig, M., & Zilibotti, F. (2017). Networks in conflict: Theory and evidence from the Great War of Africa. *Econometrica*, 85(4), 1093–132.

Konyukhovskiy, P. V. & Grigoriadis, T. (2018). *Proxy wars*. Discussion Paper 2018/4, Free University Berlin, School of Business & Economics.

Kreps, D. M. & Wilson, R. (1982). Reputation and imperfect information. *Journal of Economic Theory*, 27(2), 253–79.

Laitin, D. D. (1998). *Identity in Formation: The Russian-Speaking Populations in the Near Abroad*. Cornell University Press.

Lanoszka, A. (2016). Russian hybrid warfare and extended deterrence in Eastern Europe. *International affairs*, 92(1), 175–195.

Lee, M. M. (2018). The international politics of incomplete sovereignty: How hostile neighbors weaken the state. *International Organization*, 72(2), 283–315.

Mailath, G. T. & Samuelson, L. (2015). Reputations in repeated games. In H. P. Young & S. Zamir (Eds.), *Handbook of Game Theory with Economic Applications*, volume 4 (pp. 165–238). North Holland.

Maoz, Z. & San-Akca, B. (2012). Rivalry and state support of non-state armed groups (NAGs), 1946-2001. *International Studies Quarterly*, 56(4), 720–34.

Marshall, A. (2016). From civil war to proxy war: Past history and current dilemmas. *Small Wars & Insurgencies*, 27(2), 183–95.

Martínez, L. R. (2017). Transnational insurgents: Evidence from Colombia's FARC at the border with Chávez's Venezuela. *Journal of Development Economics*, 126, 138–53.

Maurer, T. (2015). Cyber proxies and the crisis in Ukraine. In K. Geers (Ed.), *Cyber War in Perspective: Russian Aggression against Ukraine* (pp. 79–86). NATO CCD COE publications.

McKnight, S. (2015). The rise and fall of the Rwanda-Uganda alliance (1981-1999). *African Studies Quarterly*, 15(2), 23–52.

Michalopoulos, S. & Papaioannou, E. (2016). The long-run effects of the scramble for Africa. *American Economic Review*, 106(7), 1802–48.

Morelli, M. & Rohner, D. (2015). Resource concentration and civil wars. *Journal of Development Economics*, 117, 32–47.

Mumford, A. (2013). Proxy warfare and the future of conflict. *The RUSI Journal*, 158(2), 40–6.

Padró i Miquel, G. & Yared, P. (2012). The political economy of indirect control. *The Quarterly Journal of Economics*, 127(2), 947–1015.

Pettersson, T. & Eck, K. (2018). Organized violence, 1989-2017. *Journal of Peace Research*, 55(4), 535–47.

Pieper, M. (2020). Russkiy Mir: The geopolitics of Russian compatriots abroad. *Geopolitics*, 25(3), 756–79.

Pigman, L. (2019). Russia's compatriots: Instrument or responsibility? *The RUSI Journal*, 164(2), 24–35.

Powell, R. (2013). Monopolizing violence and consolidating power. *The Quarterly Journal of Economics*, 128(2), 807–59.

Renshon, J., Dafoe, A., & Huth, P. (2018). Leader influence and reputation formation in world politics. *American Journal of Political Science*, 62(2), 325–39.

Rich, P. B. (2009). Russia as a great power. *Small Wars & Insurgencies*, 20(2), 276–99.

Robinson, J. A. (2001). Social identity, inequality and conflict. *Economics of Governance*, 2(1), 85–99.

Rohner, D., Thoenig, M., & Zilibotti, F. (2013). Seeds of distrust: Conflict in Uganda. *Journal of Economic Growth*, 18(3), 217–52.

Salehyan, I. (2010). The delegation of war to rebel organizations. *Journal of Conflict Resolution*, 54(3), 493–515.

Salehyan, I. & Gleditsch, K. S. (2006). Refugees and the spread of civil war. *International Organization*, 60(2), 335–66.

Salehyan, I., Gleditsch, K. S., & Cunningham, D. E. (2011). Explaining external support for insurgent groups. *International Organization*, 65(4), 709–44.

Salehyan, I., Siroky, D., & Wood, R. M. (2014). External rebel sponsorship and civilian abuse: A principal-agent analysis of wartime atrocities. *International Organization*, 68(3), 633–61.

Sambanis, N., Germann, M., & Schädel, A. (2018). SDM: A new data set on self-determination movements with an application to the reputational theory of conflict. *Journal of Conflict Resolution*, 62(3), 656–86.

Sambanis, N., Skaperdas, S., & Wohlforth, W. (2020). External intervention, identity, and civil war. *Comparative Political Studies*, 53(14), 2155–82.

San-Akca, B. (2016). *States in disguise: Causes of state support for rebel groups*. Oxford University Press.

Schetter, C. (2016). Playing the ethnic card: On the ethnicization of Afghan politics. *Studies in Ethnicity and Nationalism*, 16(3), 460–477.

Schultz, K. A. (2010). The enforcement problem in coercive bargaining: Interstate conflict over rebel support in civil wars. *International Organization*, 64(2), 281–312.

Siddiqi, F. H. (2012). *The politics of ethnicity in Pakistan: The Baloch, Sindhi and Mohajir ethnic movements*. Routledge.

Siddique, A. (2014). *The Pashtun question: The unresolved key to the future of Pakistan and Afghanistan*. Hurst & Company Limited.

Silve, A. & Verdier, T. (2018). A theory of regional conflict complexes. *Journal of Development Economics*, 133, 434–47.

Simonsen, S. G. (2001). Compatriot games: Explaining the 'diaspora linkage' in Russia's military withdrawal from the Baltic states. *Europe-Asia Studies*, 53(5), 771–91.

Simonsen, S. G. (2004). Ethnicising Afghanistan?: inclusion and exclusion in post-Bonn institution building. *Third World Quarterly*, 25(4), 707–729.

Smith, H. (2014). Russia as a great power: Status inconsistency and the two Chechen wars. *Communist and Post-Communist Studies*, 47(3-4), 355–63.

Spolaore, E. & Wacziarg, R. (2016). War and relatedness. *Review of Economics and Statistics*, 98(5), 925–39.

Stinnett, D. M., Tir, J., Diehl, P. F., Shafer, P., & Gochman, C. (2002). The Correlates of war (Cow) project direct contiguity data, Version 3.0. *Conflict Management and Peace Science*, 19(2), 59–68.

Timoneda, J. C. (2021). Estimating group fixed effects in panel data with a binary dependent variable: How the LPM outperforms logistic regression in rare events data. *Social Science Research*, 93, 102486.

Tremblay-Auger, B. (2019). Réputation, identités transnationales et soutien étranger de rébellions. Master's thesis, Université Laval.

van Houten, P. (1998). The role of a minority's reference state in ethnic relations. *Archives Européennes De Sociologie*, 39(1), 110–46.

Varshney, A. (2007). Ethnicity and ethnic conflict. In C. Boix & S. C. Stokes (Eds.), *The Oxford Handbook of Comparative Politics* (pp. 274–94). Oxford University Press.

Vasquez, J. A. (1995). Why do neighbors fight? proximity, interaction, or territoriality. *Journal of peace research*, 32(3), 277–293.

Vermeersch, P. (2012). Theories of ethnic mobilization: Overview and recent trends. In G. K. Brown & A. Langer (Eds.), *Elgar Handbook of Civil War and Fragile States* (pp. 223–39). Edward Elgar.

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P., & Girardin, L. (2015). Integrating data on ethnicity, geography, and conflict: The ethnic power relations data set family. *Journal of Conflict Resolution*, 59(7), 1327–42.

Walter, B. F. (2006). Building reputation: Why governments fight some separatists but not others. *American Journal of Political Science*, 50(2), 313–30.

Walter, B. F. (2009). *Reputation and Civil War: Why Separatist Conflicts are so Violent*. Cambridge University Press.

Wucherpfennig, J., Metternich, N. W., Cederman, L.-E., & Gleditsch, K. S. (2012). Ethnicity, the state, and the duration of civil war. *World Politics*, 64(1), 79–115.

Ziegler, C. E. (2006). The Russian diaspora in Central Asia: Russian Compatriots and Moscow's foreign policy. *Demokratizatsiya*, 14(1), 103–26.

Appendix

A Descriptive statistics and statuses definitions

Table 6: Descriptive statistics

	(1) All	(2) Link = 1
Triad × year-level		
Link	3.85 %	
Active support	0.31 %	2.19%
Passive support	0.44 %	2.00%
Number of countries with a link with the sponsor state	1.45	3.53
Group size in target state	12.96%	27.22%
Group size in sponsor state	3.60%	55.68%
Status in target: Monopoly or Dominant	8.71%	20.84%
Status in target: Senior or Junior Partner	19.79%	31.54%
Status in target: Powerless or Self-excluded	55.91%	32.84%
Status in target: Discriminated	15.60%	14.79%
Status in sponsor: Monopoly	1.32%	34.34%
Status in sponsor: Dominant	0.93%	24.26%
Status in sponsor: Senior Partner	1.60%	41.39%
Status in sponsor: Junior Partner, Powerless, Self-excluded, Discriminated of Irrelevant	10.40%	
Status in sponsor: Absent	85.75%	
# obs.	280,300	10,794
Sponsor × year-level		
At least one link	47.15%	
Number of links	1.26	
At least one active support	8.08%	14.02%
Number of active supports	0.10	0.19
At least one passive support	11.42%	19.77%
Number of passive supports	0.14	0.25
Number of countries with a link with the sponsor state	1.08	2.29
Group size in sponsor state	4.51%	56.45%
Status in sponsor: Monopoly	1.35%	24.78%
Status in sponsor: Dominant	1.04%	31.75%
Status in sponsor: Senior Partner	2.20%	43.46%
Status in sponsor: Junior Partner, Powerless, Self-excluded, Discriminated of Irrelevant	11.94%	
Status in sponsor: Absent	83.47%	
# obs.	8,550	4,031
Group in the target state × year-level		
At least one link	21.82%	
Number of links	0.31	
At least one active support	1.84%	2.51%
Number of active supports	0.025	0.046
At least one passive support	2.17%	1.83%
Number of passive supports	0.036	0.045
Group size in target state	17.83%	24.06%
Status in target: Monopoly or Dominant	11.86%	15.98%
Status in target: Senior or Junior Partner	26.95%	31.52%
Status in target: Powerless or Self-excluded	44.45%	36.47%
Status in target: Discriminated	16.74%	16.02%
# obs.	34,843	7,601

Statuses definitions

According to the codebook of the EPR data (Vogt et al., 2015), a group that rules alone is coded as a *Monopoly* if "Elite members hold monopoly power in the executive to the exclusion of members of all other ethnic groups"; and as *Dominant* if "Elite members of the group hold dominant power in the executive but there is some limited inclusion of "token" members of other groups who however do not have real influence on decision-making". A group that shares power is coded as a *Senior partner* if "Representatives of the group participate as senior partners in a formal or informal power-sharing arrangement. By power sharing, we mean any arrangement that divides executive power among leaders who claim to represent particular ethnic groups and who have real influence on political decision-making"; and as *Junior partner* if "Representatives participate as junior partners in government". A group is classified as *Powerless* when "Elite representatives hold no political power (or do not have influence on decision-making) at the national level of executive power - although without being explicitly discriminated against"; as *Discriminated* when "Group members are subjected to active, intentional, and targeted discrimination by the state, with the intent of excluding them from political power. Such active discrimination can be either formal or informal, but always refers to the domain of public politics (excluding discrimination in the socio-economic sphere)."; and *Self-excluded* when is has "excluded [itself] from central state power, in the sense that [it] control[s] a particular territory of the state which [it has] declared independent from the central government". Finally, groups are coded as *Irrelevant* when "they were politically relevant at the state level in a previous period. We do not record groups a) that never sought political representation at the national level or never were discriminated, b) before they gain political relevance or c) after they emigrated. In countries where ethnicity is not politicized we list the largest group as irrelevant, such as the Germans in Germany".

B Illustrative cases

In specifications 1-3, evidence of a reputation-building mechanism comes from a careful examination of the coefficient associated with the interacted term $L_{ijkt} \times N_{kt}$. From the point of view of ethnic group i in target country j, this term may vary because ethnic group i comes to (or leaves) power in sponsor country k, because another ethnic group i' comes to (or leaves) power in k, or because the set of politically relevant groups in a possible sponsor's neighborhood changes. In section 3.3, we dubbed these three sources of variation (a), (b), and (c) respectively

(a) co-ethnics enter or exit power in a possible sponsor country;

(b) other changes occur in the power coalition in a possible sponsor country;

(c) a possible sponsor country gains or loses a co-ethnic neighbor.

We take two examples to make this discussion less abstract: Afghanistan in 2005-06, and Thailand in 1979-80.

There were two ethnic groups in power in Afghanistan in 2005 according to the EPR data: the Pashtuns and the Tajiks, both coded as senior partners.[28] At that time, the Tajiks were among the politically relevant groups in Uzbekistan and Tajikistan, while the Pasthuns were politically relevant in Pakistan (Siddiqi, 2012 and Siddique, 2014). Moreover, the Persians in Iran were politically relevant, and co-ethnic with the Afghan Tajiks according to the Transborder Ethnic Kin data. In the end, according to our definition, Afghanistan as a potential sponsor state counted four co-ethnic neighbors (*i.e.* neighbors with at least one politically relevant ethnic group that is in power in Afghanistan), as it appears in Figure 2(a).

The Parliamentary elections held at the end of 2005 gave significant political influence to the Hazaras and Uzbeks (Simonsen, 2004, Ibrahimi, 2017, and Schetter, 2016). In the EPR data, this translates into these two groups becoming senior partners in 2006 – along with the Pashtuns and Tajiks. While the sets of politically relevant groups in the neighborhing countries of Afghanistan remained unchanged, the arrival of the Hazaras and Uzbeks in power implied that Turkmenistan, whose politically relevant groups included the Uzbeks, became an additional co-ethnic neighbor for Afghanistan. The country then counted five co-ethnic neighbors (Figure 2(b)).

In our data, this evolution of the ruling coalition translates into the variable N_{kt} going from 4 (at t = 2005) to 5 (at t = 2006), for k = Afghanistan. For the triads where k = Afghanistan and i = Uzbeks, this corresponds to a variation of type (a). For all other triads where k = Afghanistan, this corresponds to a variation of type (b).

As it happens, Afghanistan started to actively support two groups in Pakistan in

---

[28]Other politically relevant groups included the Hazaras and Uzbeks (both junior partners) and the Turkmens (powerless). The other ethnic groups in presence (Baloch (or Balushis), Aimaq, Nuristani, Qizalbash, Pashai, Pamir Tajiks and Brahui) were politically irrelevant according to the EPR data.

Figure 2: Afghanistan as potential sponsor state
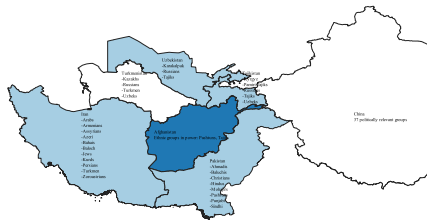
(a) 2005 : 4 co-ethnic neighbors
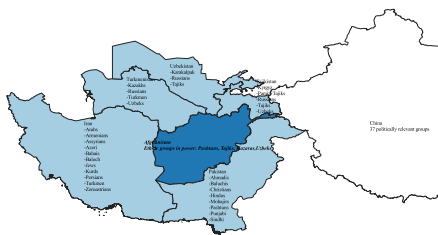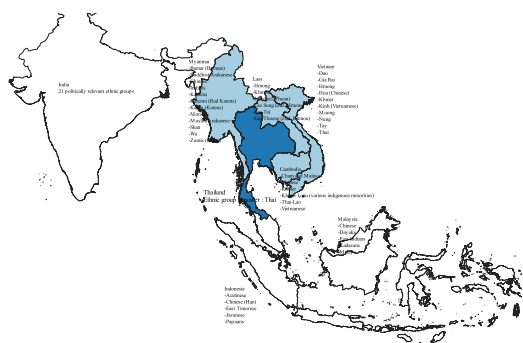

(b) 2006 : 5 co-ethnic neighbors

2008 – the Baluchis (through the *Baloch Republican Army*), who were then present but not politically relevant in Afghanistan, and the Pashtuns (through the *Tehrik-i-Taliban Pakistan* group), who were in power as senior partners. Should we interpret this timely coincidence in favor of a reputation-building mechanism? It is not clear, as new coalition members may simply mean a change in the Afghan foreign policy. Sponsor state-specific power-in-k dummies in our specification 3 allow us to eliminate variation (a), and we also control for the composition of the power coalition in sponsor states to take the best care we can of variation (b) in Table 2. In this way, we do not use the result of the 2005 Afghan Parliamentary elections to establish the existence of the reputation-building mechanism.

Figure 3 illustrates the evolution of the case of Thailand as a possible sponsor state from 1979 to 1980. In 1979, Vietnamese forces overthrew the Khmer Rouge regime in Cambodia. The new government, known as the People's Republic of Kampuchea, backed by Vietnam, focused on consolidating power among the Khmer population, and marginalized minority groups, including the Thai-Lao (Baird, 2010). In the EPR database, the Thai-Lao group lost its political relevance in Cambodia in 1980. At the time, the Thai were the dominant group in Thailand, and they were also politically relevant in Myanmar, Laos, and Vietnam. Therefore, in 1979, Thailand counted 4 co-

ethnic neighbors. When the Thai-Lao lost their political relevance in Cambodia, this number decreased to 3. For all triads where k = Thailand, the value of N_{kt} thus goes from 4 (t = 1979) to 3 (t = 1980) in our data – a change driven by a variation of type (c).

In contrast with the Afghan example discussed above, this evolution is exploited when we run our most demanding specifications with sponsor state-specific power-in-k dummies, and when we additionally control for changes in the ruling coalition in the sponsor state, because the source of variation at play here is a change in the set of politically relevant groups in the sponsor's neighborhood. Our empirical approach thus investigates whether the reduced set of co-ethnic neighbors for Thailand can plausibly be related to a lower probability that Thailand supports a rebellion from a co-ethnic group in a neighboring country (other than Cambodia), through a lower incentive to build reputation for resolve. As it happens, according to San-Akca (2016)'s data, Thailand supported several Shan groups in neighboring Myanmar from the sixties (the Shan of Myanmar being co-ethnic to the (politically dominant) Thai of Thailand, according to the Transborder Ethnic Kin data), but stopped doing so after 1984.[29]

Figure 3: Thailand as potential sponsor state

(a) 1979 : 4 co-ethnic neighbors
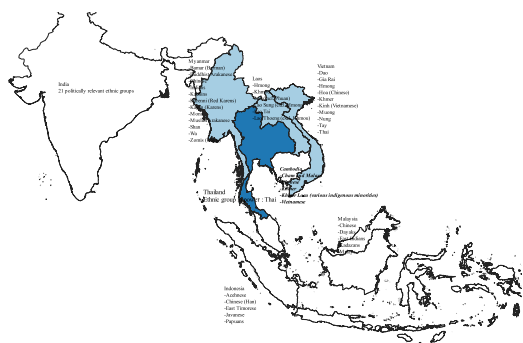
(b) 1980 : 3 co-ethnic neighbors

---

[29]In particular, an active support is reported to the insurgent *Shan State Army*, who was active from 1964 to 1973; and then to the *Shan United Revolutionary Army* and *Tai Revolutionary Council*, active from 1969 to 1984. On the contrary, Thailand is not reported to have actively supported the *Mong Tai Army* – a Shan rebel group who became active from 1985 on.

C   Additional results

a.   Weighted estimates

Table 7: Weighted estimations

	(1)	(2)	(3)	(4)	(5)	(6)
	Active support			*Political inclusion*		
Nb of countries with a link	-0.0003			-0.0039		
	(0.001)			(0.003)		
	[0.710]			[0.126]		
Link × Nb of countries with a link	0.0040	0.0074	0.0059	0.0343	0.0591	0.0503
	(0.002)	(0.003)	(0.003)	(0.020)	(0.031)	(0.038)
	[0.078]	[0.019]	[0.043]	[0.085]	[0.060]	[0.185]
Controls in Target$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes
Controls in Sponsor$_{ikt}$	Yes	Yes	Yes	Yes	Yes	Yes
Triad$_{ijk}$ FE	Yes	Yes		Yes	Yes	
Year$_t$ FE	Yes			Yes		
Target$_j$*Sponsor$_k$*Year$_t$ FE		Yes	Yes		Yes	Yes
Triad$_{ijk}$ × Power-in-k FE			Yes			Yes
R^2	0.5207	0.6843	0.6881	0.9184	0.9309	0.9319
Adjusted R^2	0.5083	0.6213	0.6258	0.9163	0.9171	0.9182

Notes: OLS estimations at the triad level. 280,300 observations (corresponding to 145 target states, 176 sponsor states, 819 groups). Controls target state include four quartile dummies for the group size among the total population in all six columns, and six political status dummies in Columns (1) to (3) (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless* and *Self-excluded*; with *Discriminated* being the group of reference). Controls sponsor state include four quartile dummies for the group size among the total population and eight political status dummies in all six columns (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless*, *Self-excluded*, *Discriminated* and *Irrelevant*, with *Absent* being the group of reference). Robust standard errors clustered at the target-country and sponsor-country levels in parentheses. P-values in squared parentheses.

b.   Country-pair level data

We collapse the data at the level of the target state × sponsor state dyad, and estimate the following model:

$$S_{jkt} = \alpha + \rho NL_{jkt} + \delta N_{kt} + \sigma NL_{jkt} \times N_{kt} + \tau_{jk} + \mu_t + \epsilon_{jkt}, \tag{6}$$

where S_{jkt} is a dummy variable equal to one if at least one ethnic group in country j is actively supported by country k at time t. NL_{jkt} is the number of ethnic links (possibly

41

greater than one) that exist between countries j and k at time t, while N_{kt} remains the number of countries with whom k shares an ethnic tie at time t. In this specification, our coefficient of interest is σ, which captures the role of the interaction between the number of ethnic links in the dyad and the number of countries with whom the sponsor state shares a co-ethnic group. We also control for dyad (τ_{jk}) and year (μ_t) fixed effects. The results are displayed in Column (1) of Table 8. In Column (2), we restrict the sample to sponsor countries that are observed in more than one dyad. We do so for comparability with Column (3), which introduces sponsor state-specific year fixed effects ξ_{kt}, and thus omits sponsor countries that belong to one dyad only.

The results point to a negative link between the number of countries that share an ethnic tie with state k and active support from k in j, possibly for dilution reasons. While the direct coefficient for the number of links in the dyad is negative, its interaction with the number of countries that share an ethnic tie with k is significantly positive, as expected in the frame of the reputation-building mechanism.

Table 8: Active support, dyad-level estimations

Dependent variable: (At least one) Active support	(1)	(2)	(3)
Nb of links in the dyad	-0.043	-0.043	-0.079
	(0.025)	(0.025)	(0.037)
	[0.084]	[0.084]	[0.030]
Nb of countries with a link with the sponsor	-0.013	-0.013	
	(0.007)	(0.007)	
	[0.061]	[0.061]	
Interaction	0.034	0.034	0.047
	(0.020)	(0.020)	(0.026)
	[0.091]	[0.090]	[0.069]
Dyad_{jk} FE	Yes	Yes	Yes
Year_t FE	Yes	Yes	
$\text{Sponsor}_k*\text{Year}_t$ FE			Yes
R^2	0.4555	0.4576	0.5617
Adjusted R^2	0.4417	0.4438	0.4325
Observations	39,693	39,152	39,152
Nb of target states	145	145	145
Nb of sponsor states	176	167	167

Notes: OLS estimations at the dyad level. Column (2): sample of sponsor countries observed in more than one dyad (for comparability with Column (3), which omits sponsor countries that belong to one dyad only as it introduces sponsor-specific year FE). Robust standard errors clustered at the target-country and sponsor-country levels in parentheses. P-values in squared parentheses.

c. Considering multiple ethnic background

Table 9: Robustness with multiple ethnic background rebel groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Active support			*Political inclusion*		
Nb of countries with a link	0.0003			-0.0050		
	(0.001)			(0.003)		
	[0.848]			[0.056]		
Link $\times$ Nb of countries with a link	0.0053	0.0084	0.0113	0.0374	0.0610	0.0735
	(0.004)	(0.005)	(0.007)	(0.020)	(0.030)	(0.038)
	[0.186]	[0.074]	[0.097]	[0.055]	[0.040]	[0.056]
Controls in Target$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes
Controls in Sponsor$_{ikt}$	Yes	Yes	Yes	Yes	Yes	Yes
Triad$_{ijk}$ FE	Yes	Yes		Yes	Yes	
Year$_t$ FE	Yes			Yes		
Target$_j$*Sponsor$_k$*Year$_t$ FE		Yes	Yes		Yes	Yes
Triad$_{ijk}$ $\times$ Power-in-k FE			Yes			Yes
R^2	0.4008	0.6093	0.6125	0.9237	0.9341	0.9349
Adjusted R^2	0.3854	0.5313	0.5350	0.9218	0.9209	0.9219

Notes: OLS estimations at the triad level. 280,300 observations (corresponding to 145 target states, 176 sponsor states, 819 groups). Controls target state include four quartile dummies for the group size among the total population in all six columns, and six political status dummies in Columns (1) to (3) (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless* and *Self-excluded*; with *Discriminated* being the group of reference). Controls sponsor state include four quartile dummies for the group size among the total population and eight political status dummies in all six columns (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless*, *Self-excluded*, *Discriminated* and *Irrelevant*, with *Absent* being the group of reference). Robust standard errors clustered at the target-country and sponsor-country levels in parentheses. P-values in squared parentheses.

d. Placebo test

Table 10: Placebo test

Dependent variable: Passive support	(1)	(2)	(3)
Nb of countries with a link	0.0007		
	(0.001)		
	[0.270]		
Link × Nb of countries with a link	0.0028	0.0023	0.0048
	(0.003)	(0.003)	(0.005)
	[0.296]	[0.422]	[0.291]
Controls in Target$_{ijt}$	Yes	Yes	Yes
Controls in Sponsor$_{ikt}$	Yes	Yes	Yes
Triad$_{ijk}$ FE	Yes	Yes	
Year$_t$ FE	Yes		
Target$_j$*Sponsor$_k$*Year$_t$ FE		Yes	Yes
Triad$_{ijk}$ × Power-in-k FE			Yes
R^2	0.4867	0.5877	0.5881
Adjusted R^2	0.4735	0.5054	0.5057

Notes: OLS estimations at the triad level. 280,300 observations (corresponding to 145 target states, 176 sponsor states, 819 groups). Controls target state include four quartile dummies for the group size among the total population and six political status dummies (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless* and *Self-excluded*; with *Discriminated* being the group of reference). Controls sponsor state include four quartile dummies for the group size among the total population and eight political status dummies (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless*, *Self-excluded*, *Discriminated* and *Irrelevant*, with *Absent* being the group of reference). Robust standard errors clustered at the target-country and sponsor-country levels in parentheses. P-values in squared parentheses.

e. Discussing the estimation method

Table 11 displays the number of observations for which the predicted outcomes fall far away from the expected interval, for each of the six specifications of Table 1. As far as active support is concerned, we find a few observations associated with a predicted value below -0.1 or above 1.1, but none below -0.5 or above 1.5, regardless of the considered specification. When we turn to political inclusion, we find a number of predicted values below -0.1 or above 2.1, with 577 to 1,981 observations across the three columns. These figures however represent a very small share of the sample: 0.2% to 0.7% of the 280,300 observations. Moreover, the predicted values remain relatively close from the expected interval, as the number of concerned observations falls drastically as soon as we restrict our attention to less conservative cutoffs.

Table 11: Number of observations with large predicted values

| Cutoffs | (1) | (2) | (3) | (4) | (5) | (6) |
		Active support			*Political inclusion*	
< −0.1; > 1.1	0	522	458			
< −0.2; > 1.2	0	72	67			
< −0.5; > 1.5	0	0	0			
< −0.1; > 2.1				577	1,959	1,981
< −0.2; > 2.2				43	494	493
< −0.5; > 2.5				0	30	30

In Table 12, we verify that our results are robust when we reiterate the specifications in Table 1 excluding observations with predicted values outside the expected interval (*i.e.*, [0,1] for Columns (1)-(3) and [0,2] for Columns (4)-(6)).

Table 12: Benchmark results excluding observations with predicted values outside the expected interval

	(1)	(2)	(3)	(4)	(5)	(6)
		Active support			*Political inclusion*	
Nb of countries with a link	-0.0065			-0.0131		
	(0.0029)			(0.0042)		
	[0.0248]			[0.0020]		
Link × Nb of countries with a link	0.0287	0.0481	0.0280	0.0681	0.1372	0.1651
	(0.0104)	(0.0147)	(0.0091)	(0.0312)	(0.0439)	(0.0596)
	[0.0058]	[0.0011]	[0.0022]	[0.0292]	[0.0018]	[0.0057]
Controls in Target$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes
Controls in Sponsor$_{ikt}$	Yes	Yes	Yes	Yes	Yes	Yes
Triad$_{ijk}$ FE	Yes	Yes		Yes	Yes	
Year$_t$ FE	Yes			Yes		
Target$_j$*Sponsor$_k$*Year$_t$ FE		Yes	Yes		Yes	Yes
Triad$_{ijk}$ × Power-in-k FE			Yes			Yes
R^2	0.4666	0.8690	0.8699	0.8896	0.9001	0.9106
Adjusted R^2	0.4385	0.8318	0.8313	0.8847	0.8721	0.8836
Observations	138,480	135,199	135,291	166,301	170,245	162,989

Notes: OLS estimations at the triad level. Columns (1)-(3) exclude observations with predicted values below 0 or above 1. Columns (4)-(6) exclude observations with predicted values below 0 or above 2. Controls target state include four quartile dummies for the group size among the total population in all six columns, and six political status dummies in Columns (1) to (3) (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless* and *Self-excluded*; with *Discriminated* being the group of reference). Controls sponsor state include four quartile dummies for the group size among the total population and eight political status dummies in all six columns (*Monopoly*, *Dominant*, *Senior partner*, *Junior Partner*, *Powerless*, *Self-excluded*, *Discriminated* and *Irrelevant*, with *Absent* being the group of reference). Robust standard errors clustered at the target-country and sponsor-country levels in parentheses. P-values in squared parentheses.