WIFO WORKING PAPERS
684/2024

Towards Rationalizing the Use of Economic Sanctions

Gabriel Felbermayr
Clifton T. Morgan
Constantinos Syropoulos
Yoto V. Yotov

Towards Rationalizing the Use of Economic Sanctions

Gabriel Felbermayr (WIFO), Clifton T. Morgan (Rice University), Constantinos Syropoulos, Yoto V. Yotov (Drexel University)

WIFO Working Papers 684/2024
August 2024

Abstract

The remarkable increase in the use of economic sanctions as a coercive tool of foreign policy over the past quarter century has been accompanied by an equally rapid growth in the number of academic and policy studies, which most often aim at quantifying the economic effects of sanctions and/or understanding the reasons for their success or failure. Interestingly, the sanctions literature has failed to answer convincingly a seemingly simple but fundamental question: Why are sanctions used in the first place? To fill this gap, we build on developments in the conflict and bargaining literatures to propose a theory that motivates the imposition of economic sanctions. We also highlight the practical relevance of our framework and discuss how it can be extended to capture additional sanction features, which are assumed away in the benchmark model.

E-Mail: gabriel.felbermayr@wifo.ac.at

2024/1/W/24024

Media owner (publisher), producer: Austrian Institute of Economic Research
1030 Vienna, Arsenal, Objekt 20 | Tel. (43 1) 798 26 01 0 | https://www.wifo.ac.at
Place of publishing and production: Vienna

Free download: https://www.wifo.ac.at/publication/pid/53727196

Towards Rationalizing the Use of Economic Sanctions

Gabriel Felbermayr
Vienna University & WIFO

T. Clifton Morgan
Rice University

Constantinos Syropoulos
Drexel University & CESifo

Yoto V. Yotov*
Drexel University & CESifo

August 15, 2023

Abstract

The remarkable increase in the use of economic sanctions as a coercive tool of foreign policy over the past quarter century has been accompanied by an equally rapid growth in the number of academic and policy studies, which most often aim at quantifying the economic effects of sanctions and/or understanding the reasons for their success or failure. Interestingly, the sanctions literature has failed to answer convincingly a seemingly simple but fundamental question: *Why are sanctions used in the first place?* To fill this gap, we build on developments in the conflict and bargaining literatures to propose a theory that motivates the imposition of economic sanctions. We also highlight the practical relevance of our framework and discuss how it can be extended to capture additional sanction features, which are assumed away in the benchmark model.

JEL Classification Codes: F13, F51, H59, N40, N70.
Keywords: Economic Sanctions, Military Conflict, Bargaining Theory.

*Acknowledgements to be added later. Contact information: Felbermayr– Professor of Economics, Vienna University of Economics and Business, Austrian Institute of Economic Research (WIFO), and CESifo. E-mail: felbermayr@wifo.ac.at; Morgan – Albert Thomas Professor of Political Science, Rice University. E-mail: morgan@rice.edu; Syropoulos – Trustee Professor of International Economics; School of Economics, Drexel Univesity, CESifo. E-mail: c.syropoulos@drexel.edu; Yotov – Professor of Economics, School of Economics, Drexel University, Research Professor, ifo Insitute, CESifo. E-mail: yotov@drexel.edu.

1 Introduction

"Sanctions are now a central tool of governments' foreign policy" (The Economist, 2021). Consistent with this observation from *The Economist*, Figure 1 depicts the evolution of existing and new sanction cases in the world between 1950 and mid-2022, based on the latest edition of the Global Sanctions Database (GSDB, Syropoulos et al., 2023).[1] The figure shows a steady increase in the number of sanction cases over time, which is particularly pronounced since the early 2000s, reaching an all-time high of more than 400 active cases by mid-2022.[2]

Unsurprisingly, the remarkable increase in the use of sanctions over the past quarter century has been accompanied by an equally rapid growth in the number and quality of academic and policy work devoted to understanding how sanctions work, to quantifying their economic effects, and to explaining the reasons for their political success or failure. Naturally (e.g., due to better data, improved econometric methods, and an urgent need for policy implications), most of the existing sanctions literature is empirical.[3] However, the changing nature of sanctions, the availability of new and better data, and new empirical findings have also revealed the need for better theory. In response, some researchers have developed new models to guide the empirical analysis (e.g., Joshi and Mahmud (2020), Morgan and Kobayashi (2021), and Ahn and Ludema (2020)), while others have adopted and applied existing theories to capture important sanction features and effects (e.g., *New Quantitative Trade Models (NQTMs)* have been used to quantify the welfare effects (Felbermayr et al., 2020b) and the extraterritorial effects (Kwon et al., 2022a) of sanctions).

Despite the significant progress that has already been made to understand sanctions and their consequences, scholars from economics and political science have struggled to convincingly answer the seemingly very simple but also fundamental question – *"Why are sanctions used in the first place?"* In part, this failure results from the fact that the recent literature has been mostly interested in quantifying the specific *effects* and consequences of sanctions rather than why they are imposed. However, the main reason appears to be that imposing sanctions is wasteful; so, rational actors should negotiate solutions. Simply said, it does not seem to make economic sense to impose sanctions. This, in turn, makes the task of motivating formally the use of sanctions a difficult one (Eaton and Engers, 1992; 1999). Thus, while we appreciate the great interest in sanctions and we value the numerous recent contributions on the topic, we believe that it is very important to fill a fundamental gap

[1] For further details on the construction and contents of the GSDB, we refer the reader to Felbermayr et al. (2020a) and Kirilakha et al. (2021). The GSDB is freely available upon request at `GSDB@drexel.edu`.

[2] Morgan et al. (2023) offers a detailed discussion of the evolution of sanctions across four eras.

[3] This is reflected well in two recent collections on economic sanctions, including Felbermayr et al. (2021) and Van Bergeijk (2021). We also refer the reader to Morgan et al. (2023) for a recent literature survey of the evolution and consequences of economic sanctions.

in our knowledge by motivating the imposition of sanctions on solid theoretical grounds. We also expect that this would lead to better empirical analysis and more informed policy actions.

Against this backdrop, our objective is twofold. First, our main goal is to develop a theory that explains why sanctions are imposed in the first place. To this end, we build on theoretical developments in the conflict and bargaining literatures in economics and political science, while capturing key aspects of sanctions. The result is a tractable theoretical model that justifies the imposition of sanctions. Naturally, in an effort to characterize the key object of interest most clearly, our benchmark model abstracts from some empirically relevant features and characteristics of the sanction process. Accordingly, our second objective is to highlight the practical implications of our theory and to demonstrate how our model can be extended to accommodate some features and characteristics of modern sanctions, which are assumed away in our initial setting.

Typically, economic sanctions are viewed as an instrument of coercive bargaining. In fact, many consider sanctions to be a substitute for the use of military force (Lopez and Cortright (1995); Mulder (2022)). Sanctions, like military force, can serve to manipulate the costs of disagreement as well as to signal the willingness of the parties to bear those costs. Thus, it is rather intuitive that we draw on bargaining theory to help us understand when sanctions are used, how they are used, and with what anticipated effect. The application of bargaining theory to the study of militarized conflict is quite well developed and two factors, in particular, suggest we should draw from existing models of conflict bargaining in our efforts to model sanctions processes. First, as indicated above, sanctions are viewed as a functional equivalent of the use of military force. Second, sanctions are similar to military force in the sense that, *ex post*, their use is inefficient Fearon (1995). That is, since they are costly to use, rational parties would be better off reaching the same outcome through bargaining. The challenge is to design a model that specifies the conditions under which sanctions are used in spite of this inefficiency.

Thus, it is informative to briefly review the literature on the emergence of militarized conflict. There are two basic modeling strategies. The currently more popular approach treats the use of force as an outside option to bargaining. Two governments involved in a dispute can either reach a negotiated agreement or they can resort to war, which will determine which side will achieve all of its aims.[4] Depending on the specific assumptions adopted, these models identify differences in risk acceptance, issue indivisibility, private information coupled with an incentive to misrepresent, or an inability of the parties to commit to abide by the terms of a negotiated settlement as a necessary condition for war.

[4]That is, if the war actually leads to a solution and not to stalemate.

(See, e.g., Fearon (1995), Powell (2004; 2006)). The second strategy is to view the use of force as an inside option in bargaining. In these models, force is used by governments in a dispute to manipulate the costs associated with continued disagreement and to signal one's tolerance for bearing such costs. These models explain war as the mechanism through which a problem of incomplete information is solved (see, Clausewitz (1968), Schelling (1966), Blainey (1973), and Wagner (2007)) and that war is avoided, or ended, when the parties agree on which outcome reflects the mutually accepted distribution of 'power'.[5]

Some advocates of the second approach have argued that war has never determined an outcome except as it influences bargaining (Keckskemeti (1958)). One can counter that the crucial factor is that war can, in principle, lead to the total destruction of one side so treating war as an outside option to bargaining is perhaps a reasonable strategy. One cannot, however, imagine any circumstances in which economic sanctions could, in fact, materially force the outcome of an interstate dispute except as they influence bargaining. For that reason, we adopt the second strategy and model sanctions as an inside option that influences bargaining.

To derive a simple and tractable model of sanctions imposition, we abstract from some elements that may be empirically and politically relevant. Therefore, to highlight the practical relevance of our analysis, in Section 3, we discuss some implications of our theory on policy-making and on empirical work. Moreover, we explain how it can be extended to accommodate additional features of sanctions that are assumed away in our benchmark setting. The first, natural implication of our theoretical analysis is that it motivates an empirical model for the imposition of sanctions, and we discuss a series of testable hypotheses. Importantly, we argue that some of the determinants of the imposition of sanctions that are implied by our theory are plausibly exogenous with respect to measures of the economic impact and political effectiveness of sanctions as key objects of interest in the empirical sanctions literature. Thus, we believe that our theory can be helpful in addressing an important but largely ignored issue in the related empirical literature – the endogeneity of sanctions.

Our theory has several direct implications for analyzing the political effectiveness of sanctions and, in particular, for explaining the puzzling finding from the literature that often sanctions do not 'work' – i.e., they do not achieve the objectives stated at their imposition. Specifically, first, our model explicitly allows for the possibility of sanctions failure, i.e., when sanctions may not achieve their political objectives and the targeted country wins the conflict. Second, our theory admits the possibility that successful sanctions may already work at the threat stage and, therefore, they may not be imposed at all. Third, our framework captures

[5]In this context, 'power' refers to more than the distribution of military capabilities and/or resources. It also reflects things like risk orientation, cost tolerance, and resolve.

the direct relationship between the costs of sanctions and the probability of their success, and it allows for improved sanction design, learning effects, and more effective enforcement that further contribute to the effectiveness of sanctions. Finally, on a related note but from a broader perspective, we hope that our interdisciplinary effort to model the imposition of sanctions may enhance our understanding of the complex sanctions processes by linking the economics literature on the economic impact/costs of sanctions and the political science literature on political effectiveness/success of sanctions.

2 A Theory of Sanctions Imposition

We develop a general, game-theoretic model, where the imposition of a 'sanction' means the application of costly mechanisms of dispute settlement as distinct from reaching a bargained agreement.[6] We use the model to show that, in equilibrium, rational players do not impose sanctions in the final period of the game but, under certain circumstances, they may use them in an earlier period. Those results are then used to identify conditions that make sanctions more likely.

2.1 The Basic Model

In our simple model, we assume the existence of two governments that interact over two periods with the second being the last. In each period, governments have two decisions, first, they decide how to allocate their resources between preparing for conflict and consumption and, second, they decide whether to settle a dispute through bargaining or to impose economic sanctions. Sanctions are costly, and thus an inefficient mechanism for dispute resolution as compared to reaching a negotiated settlement. To anticipate our conclusions, we demonstrate that sanctions never happen in the second, or ultimate, period; but, they can occur in period one if one, or both, of the governments expect to improve their bargaining position in the final period through sanctions in the first period.

We differentiate between the two governments by referring to one as the potential user or sender of sanctions, S and to the other as the potential target, T. Accordingly, we use the subscript $i \in \{S, T\}$.[7] The governments play a two-period game where the subscript

[6] Our theory is based on a broader theory of conflict resolution, some of the details of which can be found in the Appendix.

[7] In the more general version of our model, we do not use this distinction. Of course, in many cases, each government is capable of imposing sanctions, and often we observe that targets impose counter-sanctions. Our treatment is not to suggest otherwise. Rather, it is to simplify the analysis. To investigate counter-sanctions, one would reevaluate the model with the role of governments switched and with the policy in dispute being the first sender's imposition of sanctions.

$t \in \{1, 2\}$ denotes time. At the beginning of each period, governments are exogenously endowed with R_t^i of a primary resource which they can allocate between z_t^i, which is an input for the production of a generic output (or, more specifically, a consumption good), and g_t^i, which constitutes the resources devoted to preparing for economic warfare (sanctions and counter-sanctions). These allocations are subject to the resource constraint, $z_t^i + g_t^i \leq R_t^i$.

Our conceptualization of z_t^i and g_t^i merits some discussion. When applying the general model to militarized conflict, these two objects take on the familiar interpretation of butter and guns. Here, we view 'butter' as encompassing goods for consumption as well as government policy choices that impose externalities internationally. These could involve, for example, actions that result in human rights violations or environmental degradation as well as policies that involve supporting terrorists or attacking neighboring countries. As can be seen in Figure 2, these are the issues at the heart of international disputes over which sanctions are imposed.

As will be seen below, we view a portion of the goods produced by z_t^T as 'invulnerable', meaning they create no externalities and are not subject to dispute, and another portion as 'vulnerable', meaning they are subject to challenge by S. In the simplest possible setting, which we adopt here, we assume that good z_t^i can be used to produce $y_t^i = a_t^i z_t^i$ units of the consumption good, where parameter $a_t^i > 0$ captures productivity.[8] Furthermore, the representative consumer's utility coincides with y_t^i (i.e., we assume risk neutrality).

Just as a government facing the prospect of militarized conflict can choose to devote some of its resources to armaments, a government facing the prospect of economic warfare can devote resources g_t^i to preparing for that eventuality. A government S preparing for the offensive use of sanctions may want to distort the structure of comparative advantage with the objective to gain leverage over trade partners. Such distorting behavior has economic costs. For example, a government can use subsidies to develop monopoly positions on critical resources, goods or services or to acquire exclusive technological capabilities that it can weaponize. It can impose export restrictions on such goods or services or technologies in order to defend its advantages. Moreover, it must develop and support the institutional structure that will enforce those sanctions. Unlike when military force is used, sanctions cannot be simply ordered by the government. Rather, the government can only pass laws and regulations governing the economic exchanges of individuals and firms. If these laws are

[8]The analysis can be extended to consider, at the cost of analytical complexity, more general production processes. One interesting twist arises if one assumes that a is endogenous and may generate externalities. For example, a country could achieve higher a at the cost of environmental degradation or low labor standards which, in turn, create international externalities. Moreover, it is conceivable that the imposition of sanctions in period one adversely affects a country's future productivity by undermining, for example, its ability to provide high-quality education, health care, and legal services.

to be effective, an apparatus for monitoring the behavior of economic actors and enforcing those laws must be in place. Trade sanctions can only be enforced if shipped goods can be tracked, customs agents can monitor imports, and barriers against smuggling can be maintained. Financial sanctions may require fewer monitoring agents, but the technical infrastructure necessary to track trillions of transactions is considerable.

Similarly, a potential target government T may want to manipulate its country's trade structure in order to reduce strategic dependencies. In that sense, one may view g_t^i as preemptive decoupling, now often called de-risking, which has costs – foregone gains from trade – but reduces foreign leverage. Moreover, in both sender and target countries, firms and consumers facing the possibility of sanctions are likely to take actions that afford them some protection against the possibility that their business could be adversely affected. Alternative suppliers and markets may be cultivated, raw materials may be stockpiled, investment plans may be scaled back or redirected, and additional credit may be secured. Such actions can follow routine political risk analyses, but they become particularly likely once sanctions have been threatened. Many of these strategies require a direct allocation of resources, all lead to sub-optimal economic outcomes. Moreover, preparing for the possibility of sanctions incurs costs even if sanctions are never actually imposed. In this context, g_t^i represents those resources devoted to the preparation for economic warfare. We assume that the investment in g_t^i is sunk, non-contractable, and depreciates fully at the end of each time period.

Figure 3 shows that there are various types of sanction instruments ranging from classical trade sanctions that deny foreign countries access to domestic goods or services or access to domestic buyers, whether private or corporate, to financial (and investment) sanctions, travel restrictions on natural persons, military sanctions, sanctions on arms trade or other forms. Because of this multitude of types we chose to work with a very stylized model. In order to properly model one type of sanctions, say financial ones, one would need very specific modeling structures that may be very different from those required to capture other types of sanctions, say trade sanctions.

In the general model, we assume that a fraction $\kappa_t^i \in [0,1]$ of country i's output y_t^i is not contestable and that the remainder goes into a "common pool" (i.e., $\sum_j (1-\kappa_t^j) y_t^j$) which the two governments wish to claim as theirs. In short, the dispute is over this common pool and κ_t^i is an inverse measure of a government's vulnerability. For simplicity, we assume here that the sender S is invulnerable in the sense that $\kappa_t^S = 1$. The common pool thus collapses to the target's vulnerable fraction of output $(1-\kappa_t^T) y_t^T$. This captures our assumption, noted above, that economic warfare occurs over a portion of T's output, y_t^T. Note that while we assume, for tractability, that the κ_t^i's are determined exogenously, we do *not* impose the restriction $\kappa_1^i = \kappa_2^i$. Thus, it is possible that the outcome in period 1, whether through

bargaining or conflict, could influence what is in dispute in period 2. So, for example, if the target loses everything that is in dispute, whether through bargaining or conflict, in period 1, there may be nothing to dispute in period 2. In other words, security in period 2 may depend on the outcome in period 1.

In addition to determining how to allocate R_t^i, each government has a choice regarding whether to accept a bargained agreement over the distribution of $(1-\kappa_t^T)y_t^T$ or to impose sanctions over how that quantity will be distributed during that period. It is important to note that, in order to maximize its expected utility, either side can initiate conflict. In our context, this occurs when either side refuses to negotiate over the distribution of $(1-\kappa_t^T)y_t^T$ and neither accepts that the other obtains the entire surplus. Such refusal triggers the imposition of sanctions either by the sender or by the target (if it thinks it can benefit from it) or from both. The details are not important here. But what matters is that conflict is costly to both parties. Costs can be very asymmetric between T and S but, as we will introduce below, it is important to note that those costs might be quite low if, for example, the sender devotes few of its capabilities to enforcement of sanctions.

Before we describe the governments' payoff functions under each possible situation, it helps to clarify some concepts. If sanctions are imposed, the governments incur costs over and above those involved in the allocation of resources to g_t^i. Economic warfare (the imposition of sanctions) determines the distribution of $(1-\kappa_t^T)y_t^T$ for the period in which it occurs. To this end, let $\phi^i(\mathbf{g}_t|\mu_t)$, where $\mathbf{g}_t \equiv (g_t^i, g_t^j)$ and $\mu_t \equiv (\mu_t^i, \mu_t^j)$, be a sanction success function (SSF) that describes the technology of conflict;[9] that is, the probability that country i would emerge as the victor in a winner-take-all contest over the common pool in period t:

$$\phi_t^i \equiv \phi^i(\mathbf{g}_t \mid \mu_t) = \begin{cases} \frac{\mu_t^i g_t^i}{\mu_t^i g_t^i + \mu_t^j g_t^j} & \text{if} \quad g_t^i \text{ and } g_t^j > 0 \\ \frac{\mu_t^i}{\mu_t^i + \mu_t^j} & \text{if} \quad g_t^i = g^j = 0 \end{cases}, \quad i \neq j \in \{S, T\}. \tag{1}$$

One salient trait of ϕ_t^i in (1) is that it is increasing and concave in g^i and in μ^i. Another trait is that it is decreasing and convex in g^j and in μ^j. These properties ensure the existence of unique equilibria in investments of g in all the regimes we consider.[10] (When confusion is not an issue, for convenience and simplicity, we suppress to time subscript t.) Parameter μ^i (> 0) is a constant that captures how effectively an actor applies its conflict

[9]The literature on contests or tournaments uses the term conflict success function

[10]See Skaperdas (1996) for the provision of an axiomatic foundation for the form of conflict success functions (here: SSF) displayed in (1). A simpler version of this function is to assume $\mu_t^i = 1$ for $i \in \{S, T\}$, as in the pioneering contribution of Tullock (1980), to rent-seeking.

preparations in influencing its probability of winning the contest relative to its adversary.[11] Broadly speaking, we interpret this as representing a government's resolve or political will. A number of factors can influence this, one of the more important is domestic politics; i.e., the willingness of politically influential actors to bear the costs of sanctions in an effort to achieve political ends. Note that g^i and μ^i describe two different features. The former refers to investment in sanction instruments (e.g., the creation of a monopoly on rare earths), the latter to the political will to actually use those instruments in case of a conflict (e.g., the restriction of exports of rare earths).

Our aim is to capture the fact that governments often fail to implement fully their sanction policies. One can verify that $\mu^i g^i \gtreqless \mu^j g^j$ implies $\partial^2\phi^i/\partial g^i\partial\mu^i \lesseqgtr 0$ *and* $\partial^2\phi^j/\partial g^j\partial\mu^i \gtreqless 0$ for a small change in μ^i. In words, an increase in μ^i may or may not enhance either government's marginal effectiveness of its preparations–this eventuality depends on initial values. Importantly, though, when an increase in μ^i moves $\partial\phi^i/\partial g^i$ in one direction, it necessarily moves $\partial\phi^j/\partial g^j$ in the opposite direction.

Though exogenous, parameter μ^i, like κ^i, may change from period 1 to period 2. For example, perhaps due to domestic conditions related to a country's political structure, social organization, infrastructure, or domestic and international institutions, victory and/or defeat by government i in period 1 may render its efficiency in future economic warfare more or less effective. We see this as an important channel of dynamic effects. For example, victory (resp., defeat) by government i in period t may imply $\mu^i_{t+1} > \mu^i_t$ (resp., $\mu^i_{t+1} < \mu^i_t$). Moreover, parameters μ^i_t and μ^j_t could be influenced by third governments through a willingness to support one side or the other; sanctions busting would be a prominent example.

Let us now define government i's payoff function under sanctions within a particular time period t. (Discounted payoffs over the entire time horizon will be defined after payoffs under settlement within a period are defined.) In addition to diverting resources away from productive tasks through investment into economic security, actual sanctions are costly in another sense: we assume that the imposition of sanctions in period t destroys a fraction $1-\beta^i_t$ of a country's (secure and insecure) output, where $\beta^i_t \in (0,1]$. Thus, $\beta^i_t y^i_t$ is the quantity of y^i_t that survives in period t when at least one side initiates sanctions. Economic sanctions generate economic damage because they reduce the benefits from productive economic exchange, i.e., by shrinking the gains from trade. Note that the loss is country and time specific. This is done intentionally to help identify yet another channel through which sanctions may affect investment into economic security and the contenders' choices

[11] See Grossman and Kim (1995); Grossman (2001) for extensive use and justification of this parameter in various conflict theoretic settings, especially when there is a distinction between offensive and defensive armaments.

between settlement and economic warfare. It follows that the value of the contested prize is $X_t \equiv \sum_j (1-\kappa_t^j)\beta_t^j a_t^j z_t^j$. Assuming that the winner of the contest captures X_t in its entirety, we may now write country i's expected payoff function under conflict in period t as:

$$u_t^i = u_t^i(\mathbf{g}_t|\cdot) \equiv \phi_t^i X_t + \kappa_t^i \beta_t^i a_t^i z_t^i, \;\; i \in \{S, T\}. \tag{2}$$

With our special case, where the sender is secure, $X_t = (1-\kappa_t^T)\beta_t^T a_t^T z_t^T$ and

$$\begin{aligned} u_t^S &= \phi_t^S X_t + \beta_t^S a_t^S z_t^S \\ u_t^T &= \phi_t^T X_t + \kappa_t^T \beta_t^T a_t^T z_t^T. \end{aligned}$$

We now turn to peaceful divisions of the disputed prize. Let $\sigma_t^i = \sigma_t^i(\cdot)$ with $i \in \{S, T\}$ be the share of the prize that government i may obtain under a negotiated settlement in period t.[12] This share is a function that is endogenously determined depending on the bargaining protocol. For specificity, we will derive it in the context of Nash bargaining (with potentially asymmetric weights).[13] A crucial feature of bargaining here is that it is costly. Even though it averts the costs associated with conflict, governments' investments in economic security serve as leverage in their negotiations to divide the contested prize. Recall that the g_t^i are endogenously determined, irreversible, and precede all other decisions in each time period.

For given investments into economic security, negotiations in period t will aim to allocate $Y_t \equiv \sum_j (1-\kappa_t^j) a_t^j z_t^j$ between the two governments. Leaving aside for now the determination of σ_t^i, we may define government i's payoff $v_t^i(\cdot)$ under bargaining in period t as

$$v_t^i = v_t^i(\mathbf{g}_t|\cdot) \equiv \sigma_t^i Y_t + \kappa_t^i a_t^i z_t^i, \;\; i \in \{S, T\}. \tag{3}$$

2.2 Sequence of Actions

The sequence of actions in the game is as follows. At the beginning of each time period t, every government determines (simultaneously and noncooperatively) the allocation of its resource endowment R^i to economic security (preparation for economic warfare) (g_t^i) and to the production of the intermediate good (z_t^i). Policymakers then proceed to settle their dispute over the ownership of the vulnerable portion of T's output, $(1-\kappa_t^T)y_t^T$, either through bargaining or through economic warfare (the imposition of sanctions). There are additional

[12] For simplicity, we denote by σ the share of the sender and let it be constant over time.

[13] An appealing feature of Nash bargaining is that it can be viewed as the limit case of the two-player Rubinstein (1982) bargaining model that involves non-cooperative alternating offers. Since, in our case, the bargaining frontier is linear, the solution under (symmetric) Nash bargaining coincides with the solutions that would arise under several other bargaining protocols, including the Kalai-Smorodinsky, egalitarian, and equal sacrifice solutions (Anbarci et al., 2002).

costs under sanctions but not under settlement. A necessary condition for bargaining is that both sides must pursue it. Economic warfare may be initiated either unilaterally or by both sides simultaneously.

As noted above, many existing models of militarized conflict assume that conflict is the outside option in bargaining and involves an all-out war that leads to the elimination of one of the adversaries and the capture of all of its current and future output by the other side. Costly conflict is always, *ex post*, inefficient so the challenge is to develop a model that can explain its occurrence. A key feature of the international political system is that it is anarchic; that is, no authority capable of enforcing agreements exists. Explanations for militarized conflict that treat it as an outside option to bargaining lead us to conclude that one of the key factors leading to conflict is the fact that, because of anarchy, governments cannot commit credibly to abide by the terms of a negotiated settlement.

At least two problems with this common treatment exist, especially when sanctions are the mechanism of dispute resolution. First, commitment problems are always present, yet conflict is rare. Basing an explanation for conflict on the inability of governments to commit to a settlement actually makes it difficult to explain the far more common outcome. Second, conflict infrequently results in a game-ending elimination of one rival. In most cases, however, conflict culminates in (and may even pave the way to) settlement. Thus, as emphasized by Clausewitz (1968), Schelling (1966), Wagner (2007) and numerous others, it is empirically more plausible and conceptually more appealing to view conflict as part of the bargaining process, i.e., as an inside option. Clearly, in the context we are exploring here, economic sanctions *cannot* lead to the elimination of one side; therefore, we develop the model with the view that conflict is an inside option to bargaining. The challenge is to do this in a way that allows us to identify the conditions under which the governments would resort to costly measures (in this context, investment into economic security and the imposition of sanctions) to divide a valued but contested object.

We begin with the observation that all international conflicts eventually end. We might not fully understand how governments overcome commitment problems, but they seem to do so. Therefore, we assume that governments know that there will be an endgame after which the issue in dispute will be settled. As we shall develop below, we capture this in our second and final period. Our explanation for conflict, as an inside option, hinges on identifying the conditions under which a government can resort to conflict in the first period in the expectation that it will improve their prospects for a better outcome in period two.

In period 1, the disputants have two broad options. They may choose to bargain over the division of $(1-\kappa_t^T)y_t^T$ or they may choose to resort to conflict to determine their payoffs probabilistically. Crucially, for our model, whether determined by settlement or by economic

warfare, this division applies only to period 1. The actors know that neither is committed to accepting the same division in the second period, though ultimately, they may do so. The practical implication of this is that, in period 1, negotiators will know that a long-term contract on the policy issues in dispute is not feasible and will thus condition their behavior on the payoffs they expect to obtain in period 2. No less importantly, however, they will also take into account the fact that their decisions in period 1 regarding what portion of their resources to devote to preparing for conflict as well their decisions regarding conflict and settlement may have long-term consequences.

We now turn to an analysis of the game to determine under which, if any, conditions sanctions will in fact be imposed.

2.3 Analysis

In this section, we provide a solution to the game. Our objectives are to clarify the logic of the dynamic linkages in the game and to derive meaningful conditions under which sanctions dominate settlement in period 1. As usual, we follow a backward induction procedure and start with the last period. We assume that a solution to the bargaining problem is provided by the Nash bargaining model (Nash (1950) and Nash (1953)), due to its simplicity and familiarity. It should be noted, however, that in the context of the game we have specified, the bargaining protocol represented by almost any model would produce similar results.

For any given pair of investment decisions, $\mathbf{g}_2$, made at the beginning of that period, what is the division rule implied by Nash bargaining? Denote with $\lambda^S \in [0,1]$ and $\lambda^T \in [0,1]$ the sender's and the target's Nash bargaining weights, respectively, in the product $N = \left[v_2^S - u_2^S\right]^{\lambda^S} \left[v_2^T - u_2^T\right]^{\lambda^T}$, where $\lambda^S + \lambda^T = 1$.[14] Keeping in mind (from (3)) that v_2^S depends on $\sigma_2^S \in [0,1]$ (the sender's share in the prize in period 2), we can search for $\arg\max_{\sigma^S} N$. Owing to the absence of risk aversion and production complementarities, the Pareto frontier is linear. The surplus (i.e., the value of the output that would have been lost under conflict) in period 2 can be defined as

$$S_2 \equiv \textstyle\sum_j \left(1 - \beta_2^j\right) a_2^j z_2^j > 0, \tag{4}$$

Importantly, for any given investments into economic security, $\mathbf{g}_2$, such that $z_2^j > 0$, the surplus S_2 is positive. Ultimately, though, S_2 is endogenous because the investments themselves are endogenous.

With the help of the definitions of v_2^S and u_2^S in (3) and (2), respectively, one can now

[14] Note that this formulation admits the possibility that one side may capture the entire surplus (see below). This requires that we have either $\lambda^S = 1$ or $\lambda^S = 0$.

verify that, at the optimum, we will have $\sigma_2^S = (\lambda^S S_2 + u_2^S - \kappa^S a_2^S z_2^S)/Y_2$ and $\sigma_2^T = 1 - \sigma_2^S$. Substituting this solution for σ_2^i in v_2^i gives, after simplification, the following expression for value of v_2^i under Nash bargaining:

$$v_2^i = \lambda^i S_2 + u_2^i, \;\; i \in \{S, T\}. \tag{5}$$

The form of v_2^i in (5) conforms to intuition. It also reveals that, for any given (and feasible $\mathbf{g}_t$), the payoff under settlement dominates the payoff under sanctions. This is so because the existence of a positive surplus implies, *ex post*, the existence of a mutually advantageous division of the contested prize. Moreover, at an interior solution under settlement, country i's best response in terms of g_2^i, for any given g_2^j ($j \neq i$), is less than it would have been under sanctions.[15]

How does the unfolding of events in period 1 condition outcomes in period 2? Note that, if actions and events in period 1 did not affect future payoffs, then our results for period 1 would be analogous to those for period 2. In particular, settlement would dominate sanctions, so we would not observe sanctions. One possibility for sanctions to occur in period 1, is that economic warfare at that time affects the structural parameters in period 2 such that one or both actors expects to improve their payoffs in the second period by enough to justify incurring the costs of sanctions in the first. To keep the analysis simple, we assume that a negotiated settlement in period 1 does not change the parameters for period 2. Essentially, we are assuming that if the parties can resolve the commitment problem and reach a negotiated settlement in period 1, then we should expect a similar settlement in period 2.

Consider, then, what happens if conflict emerges in period 1. In this case, we must take into account two possibilities for each government i: either it *wins* the sanctions episode with probability ϕ_1^i or it *loses* the conflict with probability $1 - \phi_1^i$ ($= \phi_1^j$, $j \neq i$). The key here is to note that the structural parameters (i.e., μ_2^S and μ_2^T) under the threat of sanctions in period 2, which determine the equilibrium payoffs, v_2^S and v_2^T, may change in response to the outcome of sanctions in the preceding period. This requires us to distinguish among the various types of v_2^i payoffs that could arise in period 2. Associate "s", "w^i", and "ℓ^i" with "*settlement*", a "*win by i*", and a "*loss by i*", respectively. We may then write i's equilibrium payoff in period 2 under event $h \in \{s, w^i, \ell^i\}$ in the preceding period as $v_2^i(h)$. For example, $v_2^i(w^i)$ and $v_2^j(w^i)$ are the settlement payoffs obtained by countries i and j ($\neq i$), respectively, when country i wins (with probability ϕ_1^i in period 1).

[15] As proven in the Appendix, differentiation of (5) with respect to g_2^i and utilization of the definition of the surplus S_2 in (4) gives $\partial v_2^i / \partial g_2^i = -\lambda^i \alpha_2^i (1 - \beta_2^i) + \partial u_2^i / \partial g_2^i$, which implies that an agent will wish to dedicate a smaller quantity of resources to preparation for settlement than to preparation for conflict.

We are now in a position to define the governments' discounted payoffs under sanctions and settlement over the entire time horizon. Let $\delta \in (0,1]$ stand for countries' common discount factor. Country i's discounted payoff under sanctions in period 1 is defined as[16]

$$U^i = u_1^i + \delta \left[\phi_1^i v_2^i(w^i) + \left(1-\phi_1^i\right) v_2^i(\ell^i)\right], \;\; i \in \{S,T\}. \tag{6}$$

Next, consider country i's discounted payoff V^i under settlement. As was the case with per-period payoffs under Nash bargaining, we may write this function as $V^i = v_1^i + \delta v_2^i(s) = \sigma_1^i Y_1 + \kappa^i a_1^i z_1^i + \delta v_2^i(s)$. So, we must determine the period 1 shares (σ_1^i, σ_1^j). To this end, define

$$\Lambda(h) \equiv \textstyle\sum_j v_2^j(h), \;\; h \in \{s, w^i, \ell^i\} \tag{7a}$$

$$S_1 \equiv \textstyle\sum_j \left(1-\beta_1^j\right) a_1^j z_1^j > 0 \tag{7b}$$

$$S \equiv S_1 + \delta\left[\Lambda(s) - \phi_1^S \Lambda\left(w^S\right) - \phi_1^T \Lambda\left(w^T\right)\right]. \tag{7c}$$

For future purposes, we use the second-period payoffs in (5) to rewrite $\Lambda(h)$ in (7a), after some algebra and simplification, as

$$\Lambda(h) = a_2^S z_2^S(h) + a_2^T z_2^T(h), \;\; \text{for } h \in \{s, w^i, \ell^i\}. \tag{8}$$

The above expression clarifies that the sum of period 2 payoffs in situation h is equal to the sum of the contending countries' final goods, which depend solely on their productivities and their investments in settlement.[17]

Utilizing the fact that $\sigma_1^j = 1 - \sigma_1^i$, we can now show that

$$(\lambda^i S + U^i - \kappa^i a_1^i z_1^i)/Y_1 = \arg\max_{\sigma_1^i} \left[V^i - U^i\right]^{\lambda^i} \left[V^j - U^j\right]^{\lambda^j}.$$

Plugging the above value of σ_1^i in V^i and simplifying terms allows us to rewrite it as

$$V^i = \lambda^i S + U^i, \;\; i \in \{S,T\}, \tag{9}$$

where S is aggregate surplus. The expression for V^i in (9) may seem to suggest that, for any $\mathbf{g}_1$ in period 1, $V^i > U^i$ which suggests that both sides should also opt for settlement in

[16]Note that $\left(1-\phi_1^i\right) v_2^i(\ell^i)$ in government i's payoff in the expression below can be rewritten as $\phi_1^j v_2^i(w^j)$ for $i \neq j \in \{S,T\}$.

[17]For additional insight, consider the special case of equally productive countries, which can be captured with the normalization $a_t^i = 1$ for all i and t. Letting $R = \sum_j R^j$ and $g_2(h) = \sum_j g_2^j(h)$, we may rewrite $\Lambda(h)$ as $\Lambda(h) = R - g_2(h)$. Thus, in this case, aggregate investment into economic security is an exact measure of the sum of payoffs under situation h in period 2.

period 1. However, while this is a valid inference for period 2 when considered in isolation, a necessary condition for the conclusion to also hold true for period 1 is that the aggregate surplus S is positive. To be sure, the first component of S in (7c) is positive because settlement in period 1 generates a positive surplus $S_1 > 0$ relative to sanctions. The possibility that S may be negative arises because the second term in (7c) may be negative. This term is equal to the difference in the total payoff (output) that would be available to both sides if settlement emerged in period 1, and the weighted average of total payoffs that would arise if, instead, economic warfare emerged in period 1. The weights are the winning probabilities of the two rivals in period 1. This enables us to arrive at

Proposition 1 *A necessary condition for settlement to fail to arise as a bargaining outcome in the first period is that* $\Lambda(s) - \phi_1^S \Lambda\left(w^S\right) - \phi_1^T \Lambda\left(w^T\right) < 0$.

The above proposition suggests that, even though sanctions are costly, they may be deployed in period 1 if the inequality in Proposition 1 holds true. In words, a necessary condition for settlement to fail to arise in period 1 is that the weighted sum of the payoffs under settlement conditional on conflict arising in period 1 exceeds the payoff sum under settlement in the same period. Questions remain as to whether this possibility can – and, if so, under what conditions it will – arise.

To develop intuition on the determinants of the inequality in Proposition 1, may we start by studying the behavior of $\Lambda(s)$, evaluated at its equilibrium level, as a function of parameter μ^S. We can then compare the value of this function to the value of the weighted sum of $\Lambda\left(w^S\right)$ and $\Lambda\left(w^T\right)$ attained in period 2. As we move in the direction of addressing this issue, two points should be kept in mind. First, economic warfare is always an equilibrium (in weakly dominated strategies). This is so for the following reason. If one side decides to impose sanctions in any period, it makes no difference what the other sides does: economic warfare will be the resulting outcome and so the contenders will condition their investments on this expectation. Implicit in Proposition 1 was the assumption that communication between governments enables them to coordinate on settlement, so settlement will have a chance to arise as an equilibrium outcome if it dominates sanctions. Second, for settlement to arise as a subgame perfect equilibrium, however, it is also necessary that neither side can improve its individual payoff by deviating unilaterally from it (i.e., by imposing sanctions and adjusting its investments in economic security accordingly).

2.4 Equilibrium in the Second Period

Our argument hinges on two crucial aspects of our theory. First, parameter μ^i (and possibly κ^i, a^i, β^i, or R^i) for each contender i may change from period 1 to period 2 depending

on the outcome of sanctions. These parameters affect the value of the issue under dispute and the expected outcome, so they are critical for determining the expected utility of a conflict in period 2. This, in turn, drives the ultimate outcome associated with a negotiated settlement. This is what might allow the victor in a period 1 sanctions war to improve its expected period 2 outcome sufficiently to outweigh the costs of the sanctions. The second aspect is that the most important effect of a sanctions-induced change in the parameter(s) noted above may not be their direct effect on the expected outcome of economic warfare; rather, it may come from the indirect effect these parameters have on the amount of resources devoted to preparing for economic warfare.

Recall that, in this model, the allocation of resources to sanctions preparations, g^i, is endogenously determined. The optimal g^i is influenced by the governments' political resolve to utilize those resources when escalation occurs (i.e., μ^i), the value of the issues at stake, and the other parameters. It is possible that the most important effect of sanctions in period 1 is that it can produce a, perhaps substantial, reduction in the g's that are necessary to support the equilibrium in period 2. The clearest, partial, example of how this might work is to imagine that the victor in a period 1 sanctions episode sees an increase in its resolve for a future conflict. It could then achieve the exact same outcome that it would have expected in period 2 without the sanctions in period 1 with a lower period 2 allocation of resources to enhancing its leverage. Depending on initial values, it is possible that the period 2 equilibrium would include a reduction in investment g^i of both parties which would, in turn, increase aggregate output. So, why the governments cannot incorporate that into a period 1 agreement and save the cost of sanctions? Because the absence of an external enforcement mechanism for period 1 agreements requires agreements to be self-enforcing and the parties cannot credibly commit to doing anything in period 2 other than following their strategy that is in (subgame perfect) equilibrium.

We now turn to a sketch of the formalization of this argument with an aim to identify conditions that produce economic warfare in period 1. We proceed by first examining the equilibrium investments into economic security in period 2.[18] We then discusse how $\Lambda(\cdot)$ (the function determining the desirability of conflict in period 1) depends on the relevant parameters.

Government i's optimal level of investment into economic security is defined implicitly

[18]As already noted, our focus on settlement during period 2 is justified on the grounds that, provided sanctions are destructive in that period (i.e., $\beta_2^i < 1$ for at least one i), economic warfare is Pareto dominated by settlement and, moreover, no government has an incentive to deviate from it either (i) in stage two of the subgame by initiating sanctions, for given investments into economic security, or (ii) in stage one by adopting a different g_1^i and then proceeding to initiate sanctions while the rival sticks to peace and g_1^j accordingly.

by differentiating v_2^i in (5) with respect to g_2^i to obtain

$$\partial v_2^i/\partial g_2^i = (\partial \phi_2^i/\partial g_2^i)X_2 - c_2^i = 0, \;\; i = 1, 2. \tag{10}$$

The first term in the RHS of (10) is government i's marginal benefit (MB_2^i) to investment whereas the second term, which can be shown to be equal to

$$c_2^i = \lambda^i a_2^i (1 - \beta_2^i) + a_2^i \beta_2^i \left[\kappa^i + \left(1 - \kappa^i\right) \phi_2^i\right], \tag{11}$$

is its corresponding marginal cost (MC_2^i). Note that MB_2^i depends on the properties of the SSF described earlier and the size of the prize X_2. (Keep in mind that X_2 is endogenous and identical for the two players.) The first term in MC_2^i captures the reduction in government i's share of the surplus S_2, brought about by a marginal increase in its investment.[19] The second term captures the combined reduction in government i's secure output and the size of the contested prize X_2. For obvious reasons, MC_2^i is increasing in productivity a_2^i of i's intermediate good and its bargaining weight λ^i. It is also increasing (decreasing) in its (the rival's) investment g_2^i (g_2^j, $j \neq i$)—through their impact of the probability of winning ϕ_2^i—and increasing in the share of output that is non-disputable κ^i. Thus, all else equal, the potential sanctioner (here being the more secure player) will face a higher opportunity cost of sanctions preparations and thus will tend to invest less in those, given g_2^T. The dependence of MC_2^i on the rate of destruction $1 - \beta_2^i$ depends on how λ^i compares with $\kappa^i + (1 - \kappa^i)\,\phi_2^i$.

It is straightforward to verify that v_2^i is quasi-concave in g_2^i. This ensures the existence of continuous best-response functions and, therefore, the existence of an equilibrium in pure strategies in the subgame considered. It is also easy to check that this equilibrium is locally stable and thus unique. (See below for the derivation of this equilibrium.) Furthermore, provided the two governments' initial resource endowments are not too asymmetric, neither side will expend its entire endowment on investment into economic security.

Let a star "*" in variables identify their values at this equilibrium, normalize (without loss of generality) the μ^is in the SSF so that $\mu^i + \mu^j = 1$ ($i \neq j$) and omit the time subscript 2 to avoid cluttering. Then, keeping in mind from (11) that c^i is increasingly linearly in ϕ^i, we can describe the non-cooperative equilibrium of the period 2 problem by

$$g^{i*} = \frac{1}{c^i}\left[\frac{(\phi^{i*}\phi^{j*})\sum_n (1 - \kappa^n)\, a^n \beta^n R^n}{1 + (\phi^{i*}\phi^{j*})\sum_n (1 - \kappa^n)\, a^n \beta^n / c^n}\right], \tag{12}$$

[19] Note that negotiation increases every government i's perceived MC^i to investment by assigning to it a positive stake to the surplus—provided, of course, $\lambda^i \neq 0$. Also note that, if $\lambda^i = 0$ for both governments, there is no distinction between economic warfare and settlement, so $v_2^i = u_2^i$ and $\partial v_2^i/\partial g_2^i = \partial u_2^i/\partial g_2^i$.

where ϕ^{i*} solves

$$\Phi(\phi^i, \cdot) \equiv \frac{\mu^i \phi^j}{\mu^j \phi^i} - \frac{c^i}{c^j} = 0, \;\; i \neq j \in \{S, T\}. \tag{13}$$

The equation in (13) is derived by forming the ratios MB^i/MB^j (which from the FOCs in (10) and the definition of the SSF equals $\frac{(\partial \phi^i/\partial g^i)X_2}{(\partial \phi^j/\partial g^j)X_2} = \frac{\mu^i}{\mu^j}\frac{\mu^j g^j}{\mu^i g^i} = \frac{\mu^i \phi^j}{\mu^j \phi^i}$) and $MC^i/MC^j = c^i/c^j$. Several comments regarding (12) and (13) prove useful. First, inspection of the numerator in (12) reveals that both g^{S*} and g^{T*} rise, remain constant, or fall (by the same amount) with changes of countries' initial resource endowments that raise, preserve or reduce the sum $\sum_n (1 - \kappa^n)\, a^n \beta^n R^n$.[20] Second, from (12) one can see that g^{i*} does not depend directly on the value of agents' resolve (i.e., parameter μ^i) in the SSF. However, it does depend on μ^i indirectly through the equilibrium value of ϕ^{i*} (i.e., equilibrium power) that solves (13). In particular, from (13) and the definition of c^i in (11) one can see that, under conditions of symmetry on all parameters except μ^i, we would have $\mu^i \gtreqless \frac{1}{2} \Leftrightarrow \phi^{i*} \gtreqless \frac{1}{2}$. Third, from (12) one can verify that g^{i*} is increasing in the product $\phi^i \phi^j$ $(= \phi^i (1 - \phi^i))$ and that this product satisfies $\lim_{\phi^i \to 0} \phi^i \phi^j = \lim_{\phi^i \to 1} \phi^i \phi^j = 0$ and $\arg\max_{\phi^i} \phi^i \phi^j = \frac{1}{2}$. Moreover, because c^i (c^j) is increasing (decreasing) in ϕ^i, g^{i*} can be shown to be strictly quasi-concave in ϕ^i, and so $\arg\max_{\phi^i} g^{i*}$ exists and is unique.

With the help of the above observations we may summarize the importance of resolve μ^i for equilibrium power ϕ^{i*}, investments g^{i*} in economic security, and the sum of second-period payoffs $\Lambda(s)$ as follows:

Proposition 2 *(The Importance of Resolve) If economic warfare (sanctions) is destructive and the interacting states are symmetric in all respects – except perhaps in resolve and vulnerability – then, in the equilibrium under settlement in period 2,*

a) leverage ϕ^{i*} is increasing in resolve μ^i;

b) investment g^{i*} is hump-shaped in resolve μ^i and satisfies

(i) $\lim_{\mu^i \to 0} g^{i*} = \lim_{\mu^i \to 1} g^{i*} = 0$;

(ii) $\arg\max_{\mu^S} g^{S*} < \frac{1}{2} < \arg\max_{\mu^S} g^{T*}$;

c) the sum of payoffs $\Lambda(s)$ is U-shaped and attains a minimum at some $\mu^S > \frac{1}{2}$.

We illustrate some of the hypotheses that can follow from the above propositions with the help of Figure 4. The panels indicate how g^{i*}, ϕ^{i*}, and the sum of payoffs $\Lambda(\cdot)$ vary with changes in resolve and the amount of T's output that is vulnerable. The solid curves correspond to a hypothetical case in which $\kappa^S = \kappa^T < 1$, $\beta^S = \beta^T < 1$, $\lambda^S = \lambda^T = \frac{1}{2}$, and $a^S = a^T$, and show the importance of resolve in determining the values of our variables of

[20] Inspection of (12) also reveals that strong asymmetries in resource endowments can render one (and only one) country's resource constraint on its investment binding (which is a possibility that we abstract from here for simplicity).

interest. The dotted and dashed curves correspond to cases in which T has an increasing amount of its output at risk (i.e., κ^T rises).[21]

The blue solid-line curve in panel (b) of Figure 4 displays the relationship between government S's equilibrium leverage (or power) ϕ^{S*} and its resolve. As pointed out in part (a) of Proposition 2, this curve rises with increases in resolve μ^S. This property (and several others) can be readily inferred from (13) which requires governments' relative marginal benefits of investing in economic security to equal (in equilibrium) their corresponding relative marginal costs. To understand this relationship, first note from (13) that Φ is decreasing in ϕ^i. (This is because MB^S/MB^T is decreasing in ϕ^{S*} while MC^S/MC^T is increasing in ϕ^{S*}). Now note that an increase in S's resolve, μ^S, increases MB^S/MB^T and thus Φ. Clearly, then, to restore equilibrium, ϕ^{S*} must rise in response to an increase in μ^S.

The dependence of the equilibrium investments g^{S*} and g^{T*} of governments S and T, respectively, on resolve μ^S is captured by the blue and magenta solid-line curves in panel (a) of Figure 4. These relationships follow from the dependence of g^{S*} and g^{T*} implied by (12) and the dependence of ϕ^{S*} on resolve μ^S described above. As noted in the proposition, the two curves are hump-shaped in μ^S, with each curve converging to 0 as μ^S converges to 0 and 1, and each attaining a global maximum at $\mu^i_{\max} \in (0, \frac{1}{2})$. In words, investments in economic security (or leverage) increase for both governments as their relative resolve becomes more equal and they increase dramatically as the value of the issue in dispute increases (from the solid line to the dashed lines), especially for the target. This illustrates one of our main contentions. If sanctions in period 1 increase the disparity in resolve and reduce the value of the issues in dispute, then both parties may reduce (perhaps drastically) the resources they devote to economic security in period 2 (and thus the sum of payoffs captured by $\Lambda(\cdot)$).

Panel (c) of Figure 4 displays the sum of period 2 payoffs $\Lambda(s)$ under settlement.[22] The shape of this function can be understood by recalling (from (8)) that $\Lambda(s) = a^S z^S + a^T z^T$ and $z^i = R^i - g^i$, and by noting the dependence of g^{i*} on μ^i described in part (b) of Proposition 2. As we will see, the U shape relationship of $\Lambda(s)$ in μ^i (which is a direct consequence of the hump shape of g^{i*} in μ^i) is a cornerstone to the possible emergence of sanctions in period 1.[23] Two observations are particularly interesting. First, as T's vulnerability increases (moving from the solid line to the dashed line), the sum of the payoffs of the equilibrium bargaining settlement in period 2 decreases. This captures the intuitive conclusion that the

[21] Recall that 'resolve' refers to the willingness of governments to bear the cost of imposed sanctions and that $(1-\kappa^T_t)y^T_t$ captures the value of the issues in dispute.

[22] As explained in Proposition 1, the value of this sum (relative to the value of $\phi^S_1\Lambda(w^S)+\phi^T_1\Lambda(w^T)$) plays a key role is the determination of a positive total surplus and the desirability of sanctions in the first period.

[23] Roughly, the quasi-convexity of $\Lambda(s)$ in μ^i is important because it gives rise to the possibility that the weighted sum $\phi^S_1\Lambda(w^S)+\phi^T_1\Lambda(w^T)$, where $\Lambda(w^i)$ is associated with some $\mu^i\left(w^i\right) > \mu^i$, may exceed the $\Lambda(s)$ value implied by some μ^i that satisfies $\mu^i\left(w^j\right) < \mu^i < \mu^i\left(w^i\right)$ for $i \neq j$.

more valuable the issues at stake, the more costly economic warfare and the greater the resources devoted to for preparing it. Since more of each government's resources are devoted to economic security, the value of the outcome is decreased. Second, as T's resolve increases (relative to S's) the sum of payoffs decreases, again, because of the effect of this μ^S increases on investments. Moreover, these factors interact. We provide proof of these conclusions in the Appendix.

2.5 Shifts in Resolve

Equipped with an understanding of how resolve and vulnerability determine power and the sum of the contenders' payoffs in the second period, we may now proceed to examine the conditions that may underlie the possible emergence of sanctions in period 1. As emphasized in Proposition 1, we must identify conditions under which $\Lambda(s) < \phi_1^S \Lambda(w^S) + \phi_1^T \Lambda(w^T)$. To accomplish this task we will adopt the assumption that a "win" under sanctions by agent i in period t leads to an improvement in agent i's resolve μ^i in period $t+1$ and to no changes in it under settlement; that is, we will suppose $\mu_{t+1}^i = \mu^i\left(w_t^i\right) > \mu_t^i$ and $\mu_{t+1}^i = \mu^i\left(s_t\right) = \mu_t^i$. For these reasons, with some abuse in notation but no substantive loss, we may think of $\Lambda(\cdot)$ as $\Lambda(s) = \Lambda(\mu^i)$ and $\Lambda(w^i) = \Lambda(\mu^i\left(w^i\right))$.

We provide our argument in two steps. Step 1 involves a comparison between $\Lambda(s)$ and the average payoff $\frac{1}{2}\Lambda(w^S) + \frac{1}{2}\Lambda(w^T)$. We then turn to Step 2, which involves a comparison between $\Lambda(s)$ and $\phi_1^S \Lambda(w^S) + \phi_1^T \Lambda(w^T)$. The key to this comparison is that the probability ϕ_1^i that government i wins the contest in period 1 is allowed to depend on both governments' investments in economic security during that period. In turn, these investments depend on countries' resolve at that time. In short, the comparison hinges on the determinants of power in period 1. The substantive content and intuition behind the noted comparisons in steps 1 and 2 can be explained with the help of panels (a) and (b) of Figure 5, respectively. The figure is based on the supposition of symmetry in all aspects of the problem except in vulnerability. In particular, we assume that $\kappa^S = 1$ whereas $\kappa^T < 1$.

Step 1: As before, the horizontal axis in panel (a) captures government S's resolve μ^S under settlement in both periods. Importantly for our purposes, points $\mu^S(w^S)$ and $\mu^S(w^T)$, which are the points where the thick, green, dotted lines hit the horizontal axis, capture the possible adjustments in government S's and T's resolve under victory, respectively, in the previous period. The vertical axis measures $\Lambda(s)$ and the average payoff $\frac{1}{2}\Lambda(w^S) + \frac{1}{2}\Lambda(w^T)$ for two security values of government T: $\kappa^T = 0.6$ and $\kappa^T = 0$. The blue solid- and dotted-line curves depict $\Lambda(\mu^S)$ under these configurations of vulnerability. Focusing on the case involving $\kappa^T = 0.6$, points A and B on this curve capture the values $\Lambda(\mu^S\left(w^S\right))$ and

$\Lambda(\mu^S\left(w^T\right))$, respectively. The mean value of these payoff sums is captured by the horizontal line in magenta color. Let us now consider possible values of resolve μ^S in the range spanned by the green dotted lines, i.e., $\left[\mu^S\left(w^T\right), \mu^S\left(w^S\right)\right]$. For any initial value of μ^S in this range, we can view the endpoints of this interval as the bias in resolve, relative to μ^S, due to possible sanctions in the prior period.

Clearly, initial values of μ^S to the left of point C ensure that $\Lambda(s) < \frac{1}{2}\Lambda(w^S) + \frac{1}{2}\Lambda(w^S)$.[24] Thus, the U shape of $\Lambda(\mu^S)$ seems to tilt the balance toward satisfying the necessary condition for sanctions to arise in period 1 that was identified in Proposition 1. Notice what happens to that inequality for the various μ^S in the range under consideration when government T's property is completely vulnerable ($\kappa^T = 0$). In short, the location of μ^S in this range (and, consequently, the extent of the induced bias in resolve across the two governments) and vulnerability play key roles with regard to the possible emergence of sanctions in period 1.

For additional insight, consider the initial value of μ^S associated with the black dotted line in panel (a). Under the biases in resolve depicted by the two endpoints of the interval relative to μ^S in this case (which imply a relatively larger bias in resolve for the non-vulnerable government S) sanctions in period 1 look likely when government T's property is not too vulnerable but not necessarily when that property is very vulnerable. By the same token, if we considered initial values of μ^S to the right point C (which imply a relatively larger bias in resolve for the vulnerable government T), sanctions are more likely when government T's property is very vulnerable.

Step 2: As noted above, panel (b) captures this step. Everything in this graph is similar to the one in panel (a) except that now the relevant comparison is between $\Lambda(s)$ and $\phi_1^S\Lambda(w^S) + \phi_1^T\Lambda(w^T)$, where the latter sum depends on μ^S through the values of the winning probabilities. This explains why the solid- and dotted-line curves in magenta (that correspond to the $\kappa^T = 0.6$ and $\kappa^T = 0$ degrees of government T's vulnerability we are considering) are not horizontal lines. Of course, what is missing is an explanation of the facts that the former curve is increasing in μ^S whereas the latter curve is decreasing in μ^S. Temporarily putting aside that explanation, it is clear that the possible inferences from this figure regarding the relevant comparison of payoffs are very similar to the one in panel (a).

Interestingly, as it turns out, power ϕ_1^{i*} in period 1 under settlement and under sanctions can be shown to be increasing in μ^i. Roughly, this is so for reasons similar (but not identical) to the ones discussed in connection with ϕ_2^{i*} in Proposition 2(a) and shown in panel (b) of Figure 4 for government S. The reason the solid-line curve capturing $\phi_1^S\Lambda(\mu(w^S)) + \phi_1^T\Lambda(\mu(w^T))$ is increasing in μ^S is because ϕ_1^S (resp., ϕ_1^T) is increasing (decreasing) in μ^S. Why is this so? Because $\Lambda(\mu(w^S)) > \Lambda(\mu(w^T))$ for $\kappa^T = 0.6$ in the case considered and a

[24] Naturally, the inequality is reversed for points to the right of C.

larger resolve μ^S in period 1 increases the weight on $\Lambda(\mu(w^S))$ and reduces the weight on $\Lambda(\mu(w^T))$. Exactly the opposite is true in the case of the dotted-line curve (where $\kappa^T = 0$) because now the increases in μ^S imply a bigger (smaller) weight on the smaller $\Lambda(\mu(w^S))$ (larger $\Lambda(\mu(w^T))$.

This demonstrates that, in our model, sanctions can occur in period 1, even though they are inefficient. This follows from the dynamic nature of the model, which captures the notion that sanctions in period 1 affect resolve in period 2 and through it the equilibrium behavior in period 2. Since sanctions in period 1 can affect the resolve of the parties in period 2 (as well as possibly the value of the issues in dispute), governments can benefit sufficiently from sanctions to overcome the costs. Importantly, our model allows for either government to initiate economic warfare unilaterally, but, because this benefit of sanctions comes primarily through a reduction in the endogenously determined investment in economic security, it is possible for both governments to benefit from sanctions. It is also important to note that this result depends on our assumption that the parties cannot, in period 1, credibly commit to specific behaviors in period 2. In period 2, the opposing governments will wish to renegotiate the division of insecure output. Our task now is to discuss how this can be applied and extended to analyses of important questions regarding economic sanctions.

3 Discussion, Implications, and Extensions

To answer the question 'Why are sanctions imposed?' in a simple and tractable way, our stylized theory abstracts from some elements of the sanction process that may be important for the economic impact and political effectiveness of sanctions, e.g., the presence of third countries that may mitigate or exacerbate the negative sanction effects. In addition, even though our theory focuses on the imposition of sanctions, our model has a series of policy and empirical implications that extend beyond the imposition of sanctions per se, e.g., for quantifying the impact of sanctions on various economic outcomes.

Against this backdrop, in this section we have two related objectives. First, we wish to highlight the empirical relevance of our theory by discussing how it may affect various aspects of the sanctions process. To this end, we discuss several implications of our model. Second, we want to discuss how the model can be extended to accommodate some additional features of sanctions. Capturing all possible sanction characteristics is beyond the scope of our effort. Moreover, it may not be desirable and realistic within a single framework. Therefore, we focus on the most relevant extensions and implications of our theory.

The first, and very natural, application of our theory is that it motivates an empirical model that may be used to study the impact of various determinants of the imposition of

sanctions and the increase in their number over time. Specifically, as discussed earlier, the various comparative statics that can be performed suggest that the binary choice of whether to impose sanctions or not would depend on: (i) The value of the issues under dispute: Interestingly, the model suggests that the relationship between this and the likelihood of sanctions is not linear, or even monotonic and that it interacts with other variables (such as the actors' resolve). (ii) The governments' resolve and how resolve is affected by the imposition of sanctions: Broadly, we consider this to be determined by domestic political considerations that determine governments' willingness to bear the costs of sanctions. (iii) The governments' wealth due to superior technology and possibly abundant resource endowments and how these may be affected by imposed sanctions. We are led to expect that sanctions would typically be imposed only by the very wealthy, simply because their marginal cost of preparing for economic warfare (sanctions) is relatively lower. We also expect this to interact in interesting ways with resolve – i.e., sanctions should be most likely when a wealthy government is in dispute with a highly resolved government. (iv) The economic costs (e.g., destruction and investments in economic warfare) incurred when sanctions are actually imposed. As these costs increase, the likelihood of sanctions decreases. (v) The dynamic consequences of sanctions may extend beyond their impact on winners' resolve and, through them, on aggregate output in the future. Their impact on destruction may be long-lasting and they may affect the losing side's future productivity. Given that sanctions may be of different types, we also see a possibility to set up a discrete choice model that goes beyond the bilateral decision to sanction or not and captures the probability to impose different types of sanctions. Such a model is additionally motivated by the fact that choices regarding specific types of sanctions are influenced by the issues over which sanctions are imposed (e.g., Morgan (1995) and Morgan et al. (2023)).

A key assumption of our theory is that the imposition of sanctions is not a random process, e.g., countries/senders self-select into imposing sanctions. The implication is that sanctions are endogenous,[25] which, in turn, may lead to severe biases in the estimates of the economic impact and political success of sanctions. As discussed in Kwon et al. (2022b), the issue of sanctions endogeneity could be a serious problem in studies of the economic effects of sanctions, e.g., investigations of the impact of sanctions on trade, growth, etc. The intuition is simple and clear, for example, events that instigate sanctions – such as civil or interstate conflicts or violations of human rights – may also shape the economic effects we observe. In addition, from a political science perspective and in relation to the effectiveness of sanctions, it has been recognized that sanction success may depend on the value placed on

[25]Kaempfer and Lowenberg (2007) motivate the issue of sanction endogeneity from a political economy perspective.

the demands accompanying sanctions. However, since it is very likely that choices regarding specific sanctions are influenced by the issues over which sanctions are imposed, it might be difficult to separate the effect of sanctions from the effect of issue intractability (Morgan et al., 2023).

From an international policy perspective, the problem of sanctions endogeneity and, especially, the process of self-selecting into sanctions is reminiscent of similar issues surrounding free trade agreements (FTAs). However, unlike the international trade literature, where the potential selection bias and endogeneity of the formation of FTAs has been a central topic for a large strand of the literature,[26] much of the sanctions literature has simply ignored the potential endogeneity problem and there is no agreement on a unified and best approach for dealing with sanctions endogeneity. Some studies have tried to address the issue without IV treatment. For example, Neuenkirch and Neumeier (2015) tackle the endogenous impact of UN and US sanctions on GDP growth by reducing the control sample. Gutmann et al. (2021) employ a 'nearest neighbor matching approach' to identify the effects of sanctions on life expectancy, while Felbermayr et al. (2020b) rely on average treatment effects (ATE) methods to study the impact of sanctions on aggregate trade. Finally, in addition to relying on ATE methods, Larch et al. (2021) and Larch et al. (2022) argue that using disaggregated data (for agriculture and mining industries, respectively) may further mitigate the problem.

There are also some recent attempts to treat the endogeneity problem with standard IV methods. For example, to study the impact of sanctions on human rights, Gutmann et al. (2020) rely on instruments that are based on the target country's geographical and genetic distance from the US, and its voting alignment with the US in the UN General Assembly, while to study the sanction effects on growth, Kwon et al. (2022b) propose a general instrument based on senders' aggressiveness, which can be traced to the independent promulgation of certain laws and regulations (e.g., the US International Emergency Economic Powers Act of 1977, and the EU Maastricht Treaty of 1992). The motivation for the choice and validity of the instruments from Kwon et al. (2022b) is consistent with our theory of sanction imposition, where the creation of laws related to sanctions is part of a government's investments into economic security and the implementation of those instruments captures the sender's resolve or political will. Thus, we expect that our model and analysis of the determinants of the imposition of sanctions may offer further ideas for good instruments for sanctions endogeneity. More broadly, the econometric version of our sanctions imposition equation may be implemented as a first-stage (selection or IV) in many models that aim

[26]See for example Magee (2003), Baier and Bergstrand (2004), Baier and Bergstrand (2007), and Egger et al. (2011). Most recently, Larch and Yotov (2023) trace the evolution of the methods to evaluate the impact of trade agreements over the past 60 years.

at quantifying the impact of sanctions on various economic and political outcomes. In fact, we believe that the estimates from models that do not account for the process of imposing sanctions may be biased and, accordingly, they should be interpreted with caution.

One of the most important questions in the sanctions literature is "Do sanctions achieve their objectives?" Most of the early work on the topic concluded that sanctions were not successful in achieving their political goals, e.g., Galtung (1967), Hoffmann (1967), and Doxey (1967). The findings from those early studies may be driven by specific cases and could be subject to selection bias (i.e., they were studied exactly because they were prominent failures), more recent and more systematic investigations have reinforced the result that most sanctions are not successful. For example, Pape (1997), Hufbauer et al. (1990), and Hufbauer et al. (2007) conclude that sanctions only achieve their political objectives in about one-fourth to one-third of the cases.

Figure 6, which comes from Morgan et al. (2023) and is based on the most recent and most comprehensive sanctions data from the GSDB, reveals that the share of 'successful' sanctions has increased continuously over time. However, while there may be a positive correlation between the increased success share and the larger number of sanctions imposed, overall, sanctions are still not perceived as particularly successful policy tools and the average share of successful sanctions from Figure 6 is relatively small – about 42%. Thus, the question "Why are sanctions applied, and at an increasing rate?" remains a valid and important question, and our theory provides several answers with implications for the literature on sanction success.

First, consistent with the arguments from Morgan et al. (2021), our model captures the possibility that sanctions may reach their desired objectives already at the threat stage.[27] Thus, successful sanctions might not actually be imposed at all.[28] The empirical implication is that the stage of threats, which precedes the actual imposition of sanctions, should be taken into account when evaluating sanction success.[29] Second, consistent with some existing theoretical work, e.g., Wagner (1988) and Morgan and Schwebach (1997), our model admits

[27]While we have not modeled "threats" explicitly, the possibility of sanctions does, of course, affect outcomes. Threats are reflected in the endogenous sender's investment into economic security (and, through them, on the contenders' threat point payoffs) as well as in its exogenous resolve. Especially when the value of the disputed issue is low, the target would usually agree to a settlement that affords all of the prize to the sender.

[28]We note that some 'fake' sanctions may seem ineffective in achieving their stated objectives while they may actually be successful in achieving their true objectives, e.g., to protect domestic interests. In addition, we recognize that some sanctions are actually intended to serve symbolic Lindsay (1986) or signaling Schwebach (2000) purposes and/or to maintain the credibility of their threats Morgan et al. (2009); Bapat et al. (2013).

[29]We refer the reader to Morgan et al. (2009), Morgan et al. (2014), and Morgan et al. (2021) details on the imposition of sanction threats and for a description of the Threat and Imposition of Economic Sanctions (TIES) data base.

the possibility that sanctions may fail, i.e., that the target state may win the economic war, thus offering a direct theoretical explanation for the failure of sanctions. Third, consistent with some recent empirical evidence, our theory implies that the larger number of sanctions are imposed because their effectiveness has improved as a result of significant learning effects and improved sanction design (e.g., Early (2021) and Biersteker and Hudakova (2021)), more effective enforcement (e.g., Morgan and Bapat (2003) and Bapat and Kwon (2015)), and institutional changes that facilitate the application of sanctions (e.g., Kwon et al. (2022b)).

Importantly, our theory captures the general notion and intuitive belief that the probability of sanction success increases with the cost on the target government (Bapat et al. (2013) and Demena et al. (2021)). As described earlier, the costs of sanctions in our framework occur through 'conflict preparation' and 'output destruction'. Such general treatment of the costs of sanctions is consistent with a series of findings from the empirical literature, e.g., that governments may devote resources to shield certain economic agents that are of strategic importance, e.g., Ahn and Ludema (2020), and that sanctions are more likely to be effective when imposed on democracies than when imposed on autocracies, e.g., Allen (2008) and Lektzian and Souva (2007).

However, we also see two valuable extensions to our modeling of the costs of sanctions, which would further highlight the practical relevance of our framework. First, with respect to investment into economic security, it is possible to extend the model by introducing a rest of the world (ROW) region. This extension would allow us to explicitly capture general equilibrium (GE) diversion of economic activity toward third countries (e.g., sanctions busters or 'Black Knights,' Early (2011)), as well as the impact of extraterritorial sanctions (e.g., Kwon et al. (2022a)). An additional advantage of introducing ROW is that this would enable us to discuss the effects of multilateral sanctions, which lead to more costs for the target (e.g., Martin (1992) Bapat and Morgan (2009) and Early (2021)). Second, with respect to 'output destruction', the model can be extended to capture explicitly the direct economic links between the sender and the target, e.g., through trade, migration, FDI, etc., which may be very important for some target governments and, therefore, represent an important channel for sanction costs.

Finally, from a broader perspective, we believe that our work highlights an important interdisciplinary gap in the sanctions literature and offers a step forward in narrowing this gap. Specifically, almost all of the economics literature on sanctions has focused on their costs and economic impact. On the other hand, political scientists have focused on the political effectiveness/success of sanctions and its determinants. While it is intuitive to expect that larger economic costs may lead to an increased probability of political success and, similarly, lack of political success could be attributed to little economic impact, there is

little evidence that this is indeed the case (e.g., Bapat et al. (2013) and Demena et al. (2021)). We hope that our interdisciplinary effort to model the imposition of sanctions will enhance our understanding of the complex sanctions processes by linking the economics literature on the economic impact/costs of sanctions and the political science literature on the political effectiveness/success of sanctions.

4 Concluding Remarks

We developed a stylized model that explains why economic sanctions are used as means of foreign policy despite their obvious inefficiency. In a dynamic game, governments face commitment problems which lead them to possibly use economic sanctions in order to influence their bargaining positions in later stages of the game.

Our theory speaks to the puzzling finding from the literature that often sanctions do not 'work' by explicitly allowing for the possibility that sanctions may not achieve their political objectives and the targeted government wins the conflict. Our theory admits the possibility that successful sanctions may already work at the threat stage and, therefore, they may not be imposed at all. Our framework captures the direct relationship between the costs of sanctions and the probability of their success, and it allows for improved sanction design, learning effects, and more effective enforcement that further contribute to the effectiveness of sanctions.

Our model generates a number of testable hypotheses that could inspire future empirical work, but it also warns that functional relationships between exogenous drivers of sanctions may be non-linear and even non-monotnonic. Nonetheless, we are convinced that, by highlighting the determinants of sanctions, our theory can be useful in addressing an important but largely ignored issue in the related empirical literature – the fact that the non-random emergence of sanctions can severely bias estimates of their effects on political or economic outcomes. Our theory allows to inform two-stage procedures which, first, estimate a binary selection equation, and, second, move to estimating the effect of sanctions on economic or political outcomes. A special case of such two-stage models are instrumental variable (IV) estimates which require the existence of sanctions determinants that have explanatory power only for the imposition of sanctions (and not on outcomes) and that are plausibly exogenous to the interested outcome variable.

Finally, we hope that our work helps bridge the gap between the political science literature, which has mostly dealt with the emergence and political effects of sanctions, and scholarly work in economics, which has more strongly focused on sanctions' effects on economic outcome variables.

References

Ahn, Daniel P. and Rodney D. Ludema, "The sword and the shield: The economics of targeted sanctions," *European Economic Review*, 2020, *130* (C).

Allen, Susan Hannah, "Political Institutions and Constrained Response to Economic Sanctions," *Foreign Policy Analysis*, 2008, *4* (3), 255–74.

Anbarci, Nejat, Stergios Skaperdas, and Constantinos Syropoulos, "Comparing Bargaining Solutions in the Shadow of Conflict: How Norms against Threats Can Have Real Effects," *Journal of Economic Theory*, September 2002, *106* (1), 1–16.

Baier, Scott L. and Jeffrey H. Bergstrand, "The Economic Determinants of Free Trade Agreements," *Journal of International Economics*, 2004, *64* (1), 29–63.

_ **and** _ , "Do Free Trade Agreements Actually Increase Members' International Trade?," *Journal of International Economics*, 2007, *71* (1), 72–95.

Bapat, Navin A. and Bo Ram Kwon, "When Are Sanctions Effective? A Bargaining and Enforcement Framework," *International Organization*, 2015, *69* (1), 131–62.

_ **and T. Clifton Morgan**, "Multilateral versus Unilateral Sanctions Reconsidered: A Test Using New Data," *International Studies Quarterly*, 2009, *53* (4), 1075–94.

_ **, Tobias Heinrich, Yoshiharu Kobayashi, and T. Clifton Morgan**, "Determinants of Sanctions Effectiveness: Sensitivity Analysis Using New Data," *International Interactions*, 2013, *39* (1), 79–98.

Biersteker, Thomas J. and Zuzana Hudakova, "UN Targeted Sanctions: Historical Development and Current Challenges," *In Research Handbook on Economic Sanctions, edited by Peter A. G. van Bergeijk*, 2021, *Cheltenham, UK: Edward Elgar Publishing*, 107–124.

Blainey, Geoffrey, "The Causes of War," *New York: The Free Press*, 1973.

Demena, Binyam A., Alemayehu S. Reta, Gabriela Benalcazar Jativa, Patrick B. Kimararungu, and Peter A. G. van Bergeijk, "Publication Bias of Economic Sanctions Research: A Meta-Analysis of the Impact of Trade Linkage, Duration and Prior Relations on Sanctions Success," *In Research Handbook on Economic Sanctions, edited by Peter A. G. van Bergeijk*, 2021, *Cheltenham, UK: Edward Elgar Publishing*, 125–50.

Doxey, Margaret, "International Sanctions: A Framework for Analysis with Special Reference to the UN and Southern Africa," *International Organization*, 1967, *26* (3), 527–50.

Early, Bryan R., "Unmasking the Black Knights: Sanctions Busters and Their Effects on the Success of Economic Sanctions," *Foreign Policy Analysis*, 2011, *7* (4), 381–402.

_ , "Making Sanctions Work: Promoting Compliance, Punishing Violations, and Discouraging Sanctions Busting," *In Research Handbook on Economic Sanctions, edited by Peter A. G. van Bergeijk*, 2021, *Cheltenham, UK: Edward Elgar Publishing*, 167–86.

Eaton, J. and M. Engers, "Sanctions," *Journal of Political Economy*, 1992, *100(5)*, 899–928.

_ **and** _ , "Sanctions: Some Simple Analytics," *American Economic Review*, 1999, *89(2)*, 404–414.

Egger, Peter, Mario Larch, Kevin E. Staub, and Rainer Winkelmann, "The Trade Effects of Endogenous Preferential Trade Agreements," *American Economic Journal: Economic Policy*, 2011, *3* (3), 113–143.

Fearon, James D., "Rationalist Explanations for War," *International Organization*, 1995, *49* (3), 379–414.

Felbermayr, Gabriel, Aleksandra Kirilakha, Constantinos Syropoulos, Erdal Yalcin, and Yoto V. Yotov, "The global sanctions data base," *European Economic Review*, 2020, *129* (C).

_ **,** _ **,** _ **,** _ **, and** _ , "The global sanctions data base–Release 3: COVID–19, Russia, and multilateral sanctions," *Review of International Economics*, 2023, *June* (https://doi.org/10.1111/roie.12691).

_ **, Constantinos Syropoulos, and Yoto V. Yotov (Editors)**, "Sanctions: Theory, Quantitative Evidence and Policy Implications," Special Issue, European Economic Review, https://www.sciencedirect.com/journal/european-economic-review/special-issue/10XLM3TV9JF 2021.

_ **,** _ **, Erdal Yalcin, and Yoto Yotov**, "On the Heterogeneous Effects of Sanctions on Trade and Welfare: Evidence from the Sanctions on Iran and a New Database," School of Economics Working Paper Series 2020-4, LeBow College of Business, Drexel University May 2020.

Galtung, J., "On the Effects of International Economic Sanctions, with Examples from the Case of Rhodesia," *World Politics*, 1967, pp. 378–416.

Grossman, Herschel I., "The Creation of Effective Property Rights," *American Economic Review Papers and Proceedings*, 2001, *91* (2), 347–352.

_ **and Minseong Kim**, "Swords or Plowshares? A Theory of the Security of Claims to Property," *Journal of Political Economy*, 1995, *103* (6), 1275–1288.

Gutmann, Jerg, Matthias Neuenkirch, and Florian Neumeier, "Precision-guided or blunt? The effects of US economic sanctions on human rights," *Public Choice*, October 2020, *185* (1), 161–182.

_ **,** _ **, and** _ , "Sanctioned to Death? The Impact of Economic Sanctions on Life Expectancy and its Gender Gap," *Journal of Development Studies*, January 2021, *57* (1), 139–162.

Hoffmann, Fredrik, "The Functions of Economic Sanctions: A Comparative Analysis," *Journal of Peace Research*, 1967, *4* (2), 140–59.

Hufbauer, G. C., J. J. Schott, and K. A. Elliott, "Economic Sanctions Reconsidered: History and Current Policy," *(2nd edition). Washington, DC: Peterson Institute for International Economics.*, 1990.

_ , _ , _ , and B. Oegg, "Economic Sanctions Reconsidered," *(3rd edition). Washington, DC: Peterson Institute for International Economics.*, 2007.

Joshi, Sumit and Ahmed Saber Mahmud, "Sanctions in networks," *European Economic Review*, 2020, *130* (C).

Kaempfer, W.H. and A.D. Lowenberg, "The Political Economy of Economic Sanctions," *Handbook of Defense Economics, edited by T. Sandler and K. Hartley*, 2007, *2*.

Keckskemeti, Paul, "Strategic Surrender," *Stanford: Stanford University Press*, 1958.

Kirilakha, Aleksandra, Gabriel Felbermayr, Constantinos Syropoulos, Erdal Yalcin, and Yoto V. Yotov, "The Global Sanctions Data Base: An Update to Include the Years of the Trump Presidency," *in the Research Handbook on Economic Sanctions, Edited by Peter A.G. van Bergeijk*, 2021.

Kwon, Ohyun, Constantinos Syropoulos, and Yoto Yotov, "The Extraterritorial Effects of Sanctions," School of Economics Working Paper Series 2022-3, LeBow College of Business, Drexel University January 2022.

_ , Costas Syropoulos, and Yoto Yotov, "Do Sanctions Affect Growth?," School of Economics Working Paper Series 2022-6, LeBow College of Business, Drexel University June 2022.

Larch, Mario and Yoto Yotov, "Estimating the Effects of Trade Agreements: Lessons From 60 Years of Methods and Data," School of Economics Working Paper Series 2023-4, LeBow College of Business, Drexel University January 2023.

_ , Jeff Luckstead, and Yoto Yotov, "Economic Sanctions and Agricultural Trade," School of Economics Working Paper Series 2021-16, Drexel University 2021.

_ , Serge Shikher, Constantinos Syropoulos, and Yoto V. Yotov, "Quantifying the impact of economic sanctions on international trade in the energy and mining sectors," *Economic Inquiry*, 2022, *60* (3), 1038–1063.

Lektzian, David and Mark Souva, "An Institutional Theory of Sanctions Onset and Success," *Journal of Conflict Resolution*, 2007, *51* (6), 848–71.

Lindsay, James M., "Trade Sanctions as Policy Instruments: A Re-examination," *International Studies Quarterly*, 1986, *30* (1), 153–73.

Lopez, George A. and David Cortright, "The Sanctions Era: An Alternative to Military Intervention," *The Fletcher Forum of World Affairs*, 1995, *19* (2), 65–85.

Magee, Christopher S., "Endogenous Preferential Trade Agreements: An Empirical Analysis," *Contributions to Economic Analysis & Policy*, 2003, *2*, Articel 15.

Martin, Lisa L., "Coercive Cooperation: Explaining Multilateral Economic Sanctions," *Princeton, NJ: Princeton University Press*, 1992.

Morgan, T. Clifton, "Clinton's Chinese Puzzle: Domestic Politics and the Effectiveness of Economic Sanctions," *Issues and Studies*, 1995, *31* (8), 19–45.

_ **and Navin A. Bapat**, "Imposing Sanctions: States, Firms and Economic Coercion," *International Studies Review*, 2003, *5* (4), 65–79.

_ **and Yoshiharu Kobayashi**, "Talking to the hand: Bargaining, strategic interaction, and economic sanctions," *European Economic Review*, 2021, *134* (C).

_ **, Constantinos Syropoulos, and Yoto V. Yotov**, "Economic Sanctions: Evolution, Consequences, and Challenges," *Journal of Economic Perspectives*, Winter 2023, *37* (1), 3–30.

_ **, Navin Bapat, and Valentin Krustev**, "The Threat and Imposition of Economic Sanctions, 1971-2000," *Conflict Management and Peace Science*, 2009, *26* (1), 92–110.

_ **,** _ **, and Yoshiharu Kobayashi**, "Threat and imposition of economic sanctions 1945-2005: Updating the TIES dataset," *Conflict Management and Peace Science*, 2014, *31* (5), 541–558.

_ **,** _ **, and** _ , "The Threat and Imposition of Economic Sanctions Data Project: A Retrospectivet," *In Research Handbook on Economic Sanctions, edited by Peter A.G. van Bergeijk*, 2021, *Chetenham, UK: Edward Elgar Publishing*, 44–61.

Morgan, T.C. and V. L. Schwebach, "Fools Suffer Gladly: The Use of Economic Sanctions in International Crises," *International Studies Quarterly*, 1997, *41*, 27–50.

Mulder, Nicholas, "The Economic Weapon: The Rise of Sanctions as a Tool of Modern War," *New Haven: Yale University Press*, 2022.

Nash, John F., "The Bargaining Problem," *Econometrica*, 1950, *18*, 155–162.

_ , "Two-Person Cooperative Games," *Econometrica*, 1953, *21*, 128–140.

Neuenkirch, Matthias and Florian Neumeier, "The impact of UN and US economic sanctions on GDP growth," *European Journal of Political Economy*, 2015, *40* (PA), 110–125.

Pape, R.A., "Why Economic Sanctions Do Not Work," *International Security*, 1997, *22(2)*, 90–136.

Powell, Robert, "The Inefficient Use of Power: Costly Conflict with Incomplete Information," *American Political Science Review*, 2004, *98* (2), 231–241.

_ , "War as a Commitment Problem," *International Organization*, 2006, *60* (1), 169–203.

Rubinstein, Ariel, "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 1982, *50* (1), 97–109.

Schelling, Thomas C., "The Strategy of Conflict," *Cambridge: Harvard University Press*, 1966.

Schwebach, Valerie L., "Sanctions as Signals: A Line in the Sand or a Lack of Resolve?," *In Sanctions as Economic Statecraft, edited by Steve Chan and A. Cooper Drury, 187–211*, 2000, *International Political Economy Series. London: Palgrave Macmillan.*

Skaperdas, Stergios, "Contest success functions (*)," *Economic Theory*, 1996, *7* (2), 283–290.

The Economist, "Sanctions are now a central tool of governments' foreign policy," Technical Report, https://www.economist.com/finance-and-economics/2021/04/22/sanctions-are-now-a-central-tool-of-governments-foreign-policy 2021.

Tullock, G., "Efficient Rent Seeking," *In: Buchanan, J., Tollison, R. and Tullock, G., Eds., Toward a Theory of Rent Seeking Society*, 1980, *Texas A and M University Press, College Station*, 97–112.

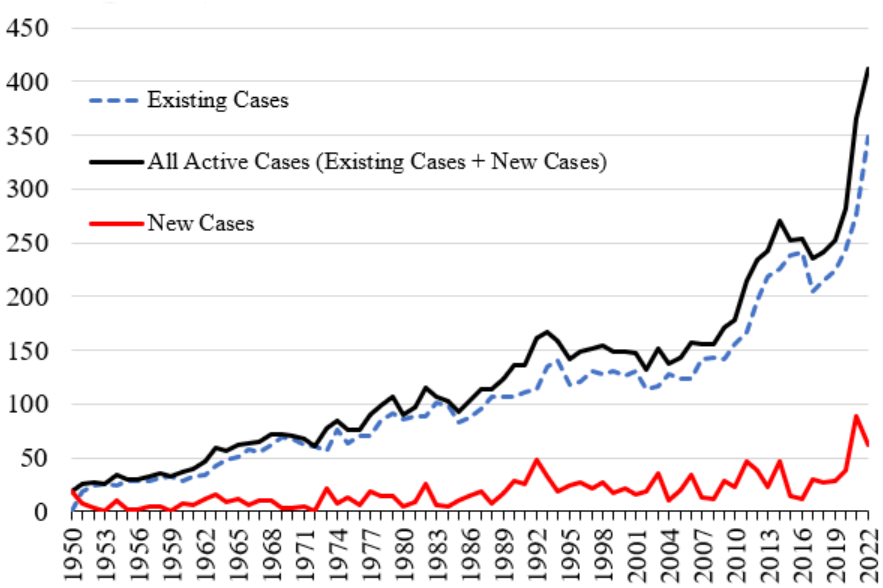
Van Bergeijk, Peter A. G. (Editor), *Research Handbook on Economic Sanctions*, Edward Elgar Publishing Ltd, 2021.

von Clausewitz, Carl, "On War. Ed. Anatol Rapoport," *New York: Penguin Press*, 1968.

Wagner, R. Harrison, "Economic Interdependence, Bargaining Power, and Political Influence," *International Organization*, 1988, *42* (3), 461–83.

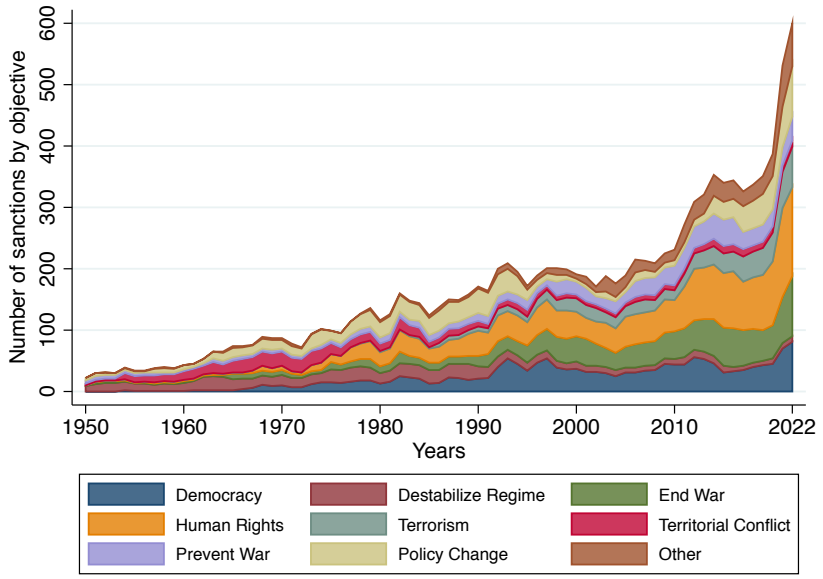
_ , "War and the State," *Ann Arbor: University of Michigan Press*, 2007.

Figure 1: Evolution of economic sanctions, 1950-2022



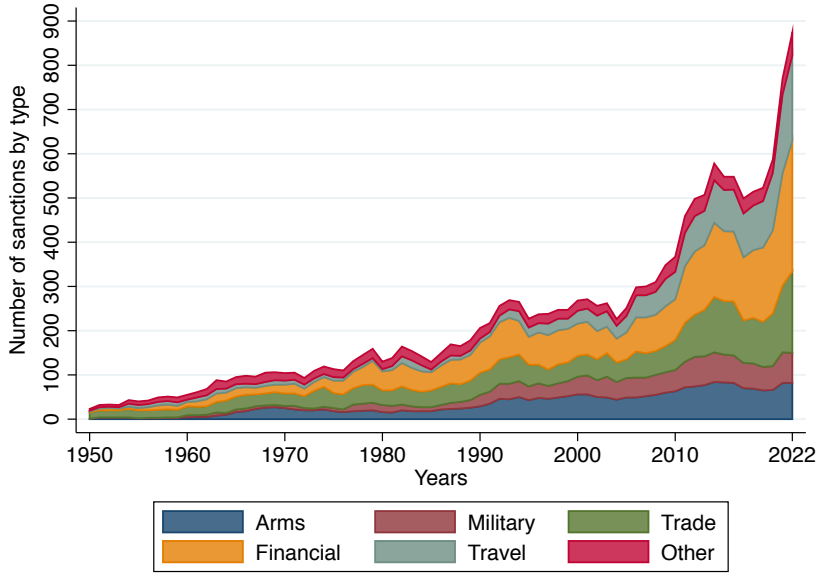
Note: This figure, which appears as panel (a) of Figure 1 in Syropoulos et al. (2023), illustrates the number of all active sanctions (dark solid line), all pre-existing (excluding terminated) sanctions (blue dashed line), and newly imposed sanctions (red solid line) in each year between 1950 and mid-2022. Additional details are provided in the main text of Syropoulos et al. (2023).

Figure 2: Evolution of Sanctions by Objective, 1950–2022



Note: This figure is from Morgan et al. (2023), where it appears as Figure 3. The figure displays the evolution of sanctions depending on their objective over the period 1950–2022. The coverage stops before the middle of 2022. The range on the y-axis of this figure is longer as compared to the range in Figure 1 because some sanction cases include more than one type of sanction. The original data used to construct the figure is the third release of the Global Sanctions Data Base (GSDB), and we refer the reader to Felbermayr et al. (2020a) and Felbermayr et al. (2023) for definitions and examples for the different sanction objectives, and details on the construction of the GSDB.

Figure 3: Evolution of Sanctions by Targeted Impact, 1950–2022



Note: This figure is from Morgan et al. (2023), where it appears as Figure 2. The figure displays the evolution of sanctions depending on their type over the period 1950–2022. The coverage stops before the middle of 2022. The range on the y-axis of this figure is longer as compared to the range in Figure 1 because some sanction cases include more than one type of sanction. The original data used to construct the figure is the third release of the Global Sanctions Data Base (GSDB), and we refer the reader to Felbermayr et al. (2020a) and Felbermayr et al. (2023) for definitions and examples for the different types of sanctions, and details on the construction of the GSDB.

Figure 4: The Effects of Resolve and Insecurity in the Second Period

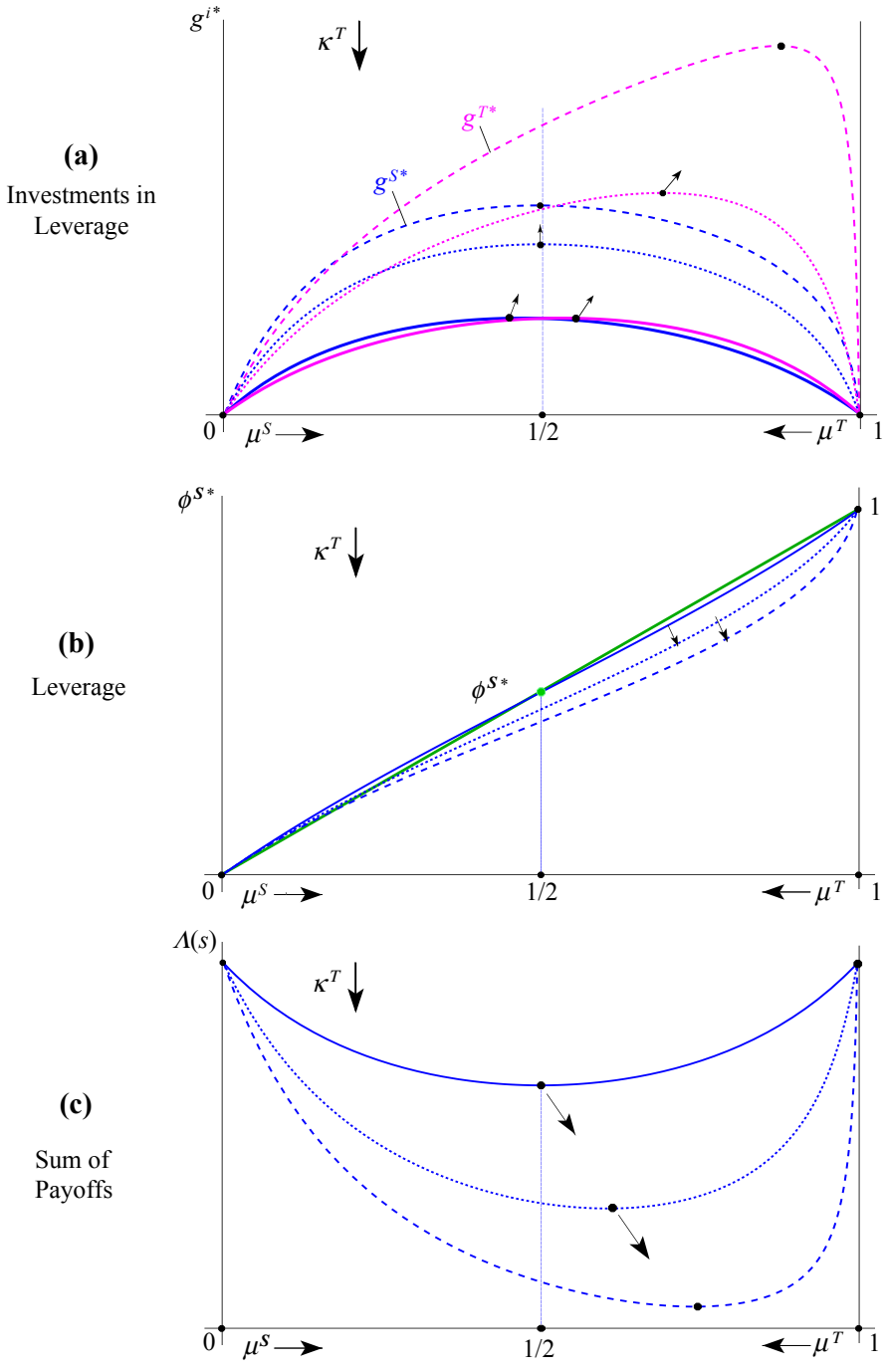


Figure 5: Resolve and Incentives to Initiate Conflict in the First Period

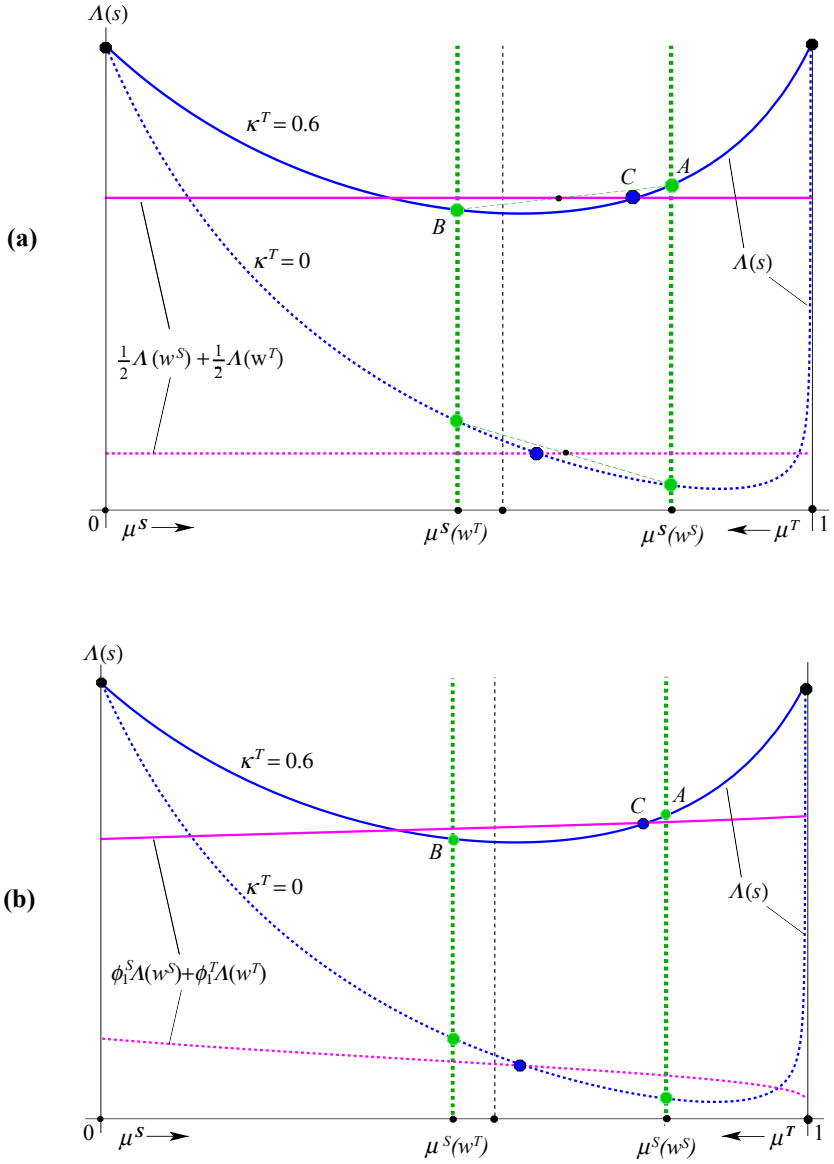
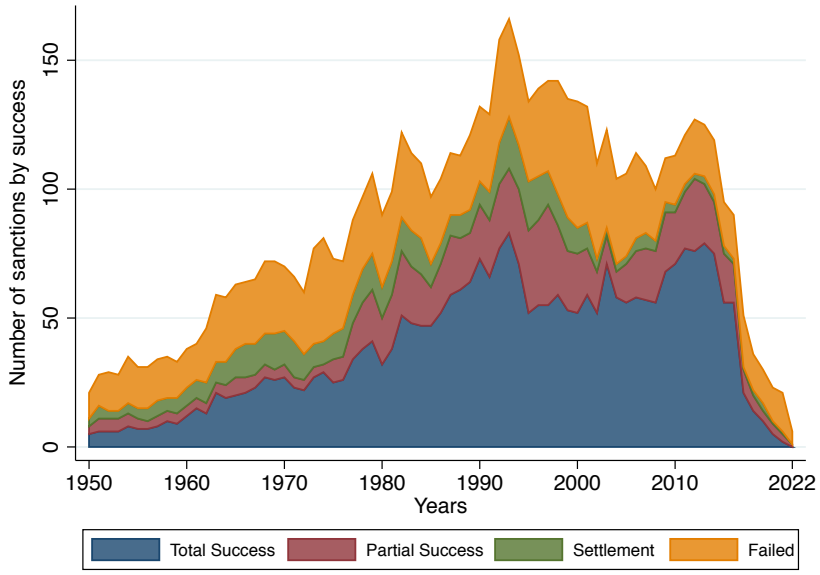


Figure 6: Evolution of Sanctions by Success, 1950–2022



Note: This figure is from Morgan et al. (2023), where it appears as Figure 4. The figure displays the evolution of sanctions depending on the success of reaching their individual political objectives, 1950–2022. The coverage stops before the middle of 2022. Since some sanctions include more than one objective, success is defined for each individual objective. Ongoing sanctions are not included in the data used to construct these figures. The original data used to construct the figure is the third release of the Global Sanctions Data Base (GSDB), and we refer the reader to Felbermayr et al. (2020a) and Felbermayr et al. (2023) for definitions, examples, and details on the construction of the GSDB.

Appendix

Our primary goal in this appendix is to shed more light on our theory and to substantiate some of our key findings. We will do this by relabeling the interacting governments as 1 and 2, and by not imposing any restrictions on their initial levels of vulnerability; that is, we let $\kappa^i \in [0,1)$ for $i = 1, 2$.

We start with the claim in footnote #15 that, for any given g^j, government i will have an incentive to invest less in preparation for settlement than in preparation for economic warfare in period 2. From (5), we have $\partial v^i/\partial g^i = \lambda^i \left(\partial S_2/\partial g^i\right) + \partial u^i/\partial g^i$, where $\partial S_2/\partial g^i = -\left(1-\beta^i\right) a^i < 0$. Moreover, the FOC for an interior solution to country i's optimal investment under conflict is given by the solution to $\partial u^i/\partial g^i = 0$. (It is very easy to show that, provided the resource constraint is not binding, this solution exists and is unique.) Denote this solution with g_c^{i*} and suppose that under settlement country i chooses $g^i = g_c^{i*}$. It follows that $\partial v^i/\partial g^i\big|_{g^i = g_c^{i*}} = -\lambda^i \left(1-\beta^i\right) a^i \leq 0$, which implies that government i will have an incentive to reduce its investment g^i below g_c^{i*}; therefore, $g_s^{i*} < g_c^{i*}$ for any given and feasible g^j, as claimed in footnote #15.

Proof of Proposition 2: Starting with the definition of $\Phi(\cdot)$ in (13), for convenience and ease of interpretation, it helps to rewrite the ratio of marginal benefits and the ratio of marginal costs of investing in leverage as

$$B = B(\phi^i, \mu^i) \equiv \frac{\mu^i \phi^j}{\mu^j \phi^i} \quad (= MB^i/MB^j)$$
$$C = C(\phi^i, \cdot) \equiv \frac{c^i}{c^j} \qquad (= MC^i/MC^j),$$

respectively, where (from (11)) we have

$$c^i = \lambda^i a^i \left(1-\beta^i\right) + a^i \beta^i [\kappa^i + (1-\kappa^i)\phi^i],$$

to obtain

$$\Phi(\cdot) = B(\phi^i, \mu^i) - C(\phi^i, \cdot). \tag{A.1}$$

Part (a): Utilizing the facts that $\phi^j = 1 - \phi^i$ and $\mu^j = 1 - \mu^i$, differentiation of the benefit schedule B gives $B_{\phi^i} = -\frac{\mu^i/\mu^j}{(\phi^i)^2} < 0$ and $B_{\phi^i\phi^i} = \frac{2\mu^i/\mu^j}{(\phi^i)^3} > 0$; therefore, B is decreasing and convex in ϕ^i. Moreover, $B \to \infty$ as $\phi^i \to 0$ while $B \to 0$ as $\phi^i \to 1$. Turning to the cost schedule C, one can verify the following: $C_{\phi^i} = C\left(\eta^i + \eta^j\right) > 0$ for $i \neq j$ because $\eta^i = \frac{\left(1-\kappa^i\right)\beta^i}{\lambda^i(1-\beta^i)+\kappa^i\beta^i+(1-\kappa^i)\beta^i\phi^i} > 0$; $C_{\phi^i\phi^i} > 0$; and, lastly, $\lim_{\phi^i \to 1} C > \lim_{\phi^i \to 0} C > 0$.

From the above, we have $\Phi_{\phi^i} = B_{\phi^i} - C_{\phi^i} < 0$, so Φ is decreasing in ϕ^i.[30] Moreover, $\lim_{\phi^i \to 0} \Phi = \infty$ while $\lim_{\phi^i \to 1} \Phi < 0$. It follows that there is a unique solution ϕ^{i*} to $\Phi(\phi^i, \cdot) = 0$. This solution can be visualized with the help of the benefit and cost schedules B and C, respectively, that are illustrated by the solid-line curves in Figure 7. It is easy for one to check that the equilibrium value $\phi^{i*} = \frac{1}{2}$ in the figure arises under the assumption of symmetry in the corresponding parameters of the two players, including $\mu^i = \frac{1}{2}$ (symmetric resolve).

Standard comparative statics exercises with respect to the various parameters of the model are now possible. For example, the impact of resolve μ^i on Φ is given by $\Phi_{\mu^i} = B_{\mu^i} = \frac{\phi^j/\phi^i}{(\mu^j)^2} > 0$. Therefore, $d\phi^{i*}/d\mu^i = -\Phi_{\mu^i}/\Phi_{\phi^i} > 0$. This adjustment in ϕ^{i*} is shown in Figure 7 and it arises because of the shift in the benefit schedule from the blue, solid-line curve to the blue, dotted-line curve. In words, at any $\phi^{i*} > 0$, a marginal increase in government i's relative resolve ($\mu^i \uparrow$) brings about an increase in its leverage. That is precisely what panel (b) of Figure 4 captures for $i = S$ and $j = T$.

One can also verify that $\Phi_{\kappa^j} = -C_{\kappa^j} > 0$. Since $d\phi^{i*}/d\kappa^j = -\Phi_{\kappa^j}/\Phi_{\phi^i} > 0$ in this case, an increase in government j's vulnerability ($\kappa^j \downarrow$) induces it to invest more heavily in economic security relative to its rival i, so i's leverage falls ($\phi^{i*} \downarrow$). This point is also captured by the shifted lines in panel (b) of Figure 4. This completes the proof of part (a).

Part (b): To prove this part, we rely on (12) to rewrite g^{i*} as

$$g^{i*} = \frac{FH}{c^i + F\Theta}, \tag{A.2}$$

where

$$F \equiv \phi^i \phi^j, \quad i \neq j = 1, 2, \tag{A.3a}$$

$$H \equiv \textstyle\sum_n (1 - \kappa^n)\, a^n \beta^n R^n, \tag{A.3b}$$

$$\Theta \equiv \left(1 - \kappa^i\right) a^i \beta^i + C \left(1 - \kappa^j\right) a^j \beta^j. \tag{A.3c}$$

Now observe from (A.3a) that $\lim_{\phi^i \to 0} F = 0 = \lim_{\phi^i \to 1} F$, along with part (a) of the proposition, implies $\lim_{\mu^i \to 0} F = \lim_{\mu^i \to 1} F = 0$ and thus $\lim_{\mu^i \to 0} g^{i*} = 0 = \lim_{\mu^i \to 1} g^{i*}$.

Next, recognizing the dependence of c^i and C on ϕ^i, note that $F_{\phi^i} = \phi^j - \phi^i$ and $\Theta_{\phi^i} = C_{\phi^i} (1 - \kappa^j)\, a^j \beta^j > 0$ (since $C_{\phi^i} > 0$ from part (a)). Differentiating g^{i*} in (A.2) with respect

[30] The sign of $\Phi_{\phi^i \phi^i}$ can be negative or positive (so Φ may be concave or convex) in ϕ^i, depending on parameter values. One can show that, under conditions of symmetry on all parameters except μ^i, we have $\Phi_{\phi^i \phi^i} \lesseqgtr 0$ as $\mu^i \lesseqgtr \frac{1}{2}$.

to ϕ^i gives, after some rearrangement and simplification,

$$\frac{dg^{i*}}{d\phi^i} = -\frac{H}{\left(c^i + F\Theta\right)^2}\left[\left(\phi^i - \phi^j\right)c^i + F\left(1-\kappa^i\right)\beta^i a^i\left(1 + F\left(1-\kappa^j\right)\beta^j a^j/c^j\right)\Omega\right], \quad \text{(A.4)}$$

where

$$\Omega = 1 + \frac{Fc^i\left[\left(1-\kappa^j\right)\beta^j a^j/c^j\right]^2}{\left(1-\kappa^i\right)\beta^i a^i\left[1 + F\left(1-\kappa^j\right)\beta^j a^j/c^j\right]} > 0. \quad \text{(A.5)}$$

Evaluating $dg^{i*}/d\phi^i$ in (A.4) at $\phi^i = 0$ and $\phi^i \geq \frac{1}{2}$ ($\Leftrightarrow \phi^j - \phi^i \leq 0$) readily implies

$$dg^{i*}/d\phi^i\big|_{\phi^i=0} = \frac{H/a^i}{\lambda^i\left(1-\beta^i\right) + \kappa^i\beta^i} > 0 \quad \text{(A.6a)}$$

$$dg^{i*}/d\phi^i\big|_{\phi^i\geq 1/2} < 0. \quad \text{(A.6b)}$$

The above inequalities establish that g^{i*} rises in ϕ^i and thus by part (a) in resolve μ^i for μ^i sufficiently close to 0. However, g^{i*} necessarily falls when ϕ^i becomes sufficiently large which, again by part (a), happens when resolve μ^i is large enough. Thus, g^{i*} is hump-shaped.

Suppose now that the governments of the interacting states are symmetric in all aspects except perhaps in resolve. Since in this case $\phi^i - \phi^j \gtreqless 0$ as $\mu^i \gtreqless \frac{1}{2}$, it follows from (A.4) that g^{i*} reaches its peak at some $\mu^i < \frac{1}{2}$. Turning to the case in which country i is more vulnerable than country j (i.e., $\kappa^i > \kappa^j$), while retaining symmetry in all other parameters (except perhaps μ^i), with some additional work we can still prove the following: (i) $\arg\max_{\mu^i} g^{i*} < \frac{1}{2}$; (ii) $\arg\max_{\mu^i} g^{i*} > \arg\max_{\mu^j} g^{j*}$; and (iii) $\arg\max_{\mu^i}\left(g^{i*} + g^{i*}\right) > \frac{1}{2}$.

Part (c): This part can be proven with the help of parts (a) and (b). ||

Next, suppose the interacting governments choose settlement in period 1. We can show that a unique equilibrium (g_1^{i*}, g_1^{j*}) exists and we can characterize its dependence on the various parameters of the model, including policymakers' resolve, the states' vulnerabilities, etc. In fact, our analysis in step 2 in the main body of the paper was based on this analysis. Once this equilibrium is obtained, we can set investments in period 1 at (g_1^{i*}, g_1^{j*}) and then ask: would any policymaker wish to deviate from settlement in period 1 by initiating economic warfare? Our analysis in the text supplies the answer. Specifically, we argued that settlement may fail to be incentive compatible if conflict in period 1 reinforces a winner's resolve so that aggregate arming in period 2 falls enough to outweigh the cost of conflict in period 1. In short, bargaining in period 1 may fail to deliver settlement at that time.

We can then ask another question: If a policymaker plans to deviate from settlement in period 1 by declaring initiating conflict, then that policymaker will also have an incentive to adjust it arming optimally in the same period and then proceed to declare economic warfare.

Since in this case the policymaker is endowed with additional flexibility, it becomes even more likely that settlement will be undermined by such a unilateral deviation in period 1. Thus, our finding that bargaining in period 1 may fail is strengthened.

Figure 7: Cost vs. Benefit Schedules and Analysis