TARGETING AUTOCRATS: ECONOMIC SANCTIONS AND REGIME CHANGE*

Manuel Oechslin,

University of Bern / World Trade Institute†

February 21, 2011

Abstract

When it comes to international economic sanctions, the most frequent goal is *regime change and democratization.* Up to now, however, such sanctions have usually failed to achieve their stated goal. Paradoxically, in some cases (e.g., Haiti, Iraq), they even made the targeted regimes resort to policies which severely amplified the direct negative economic consequences. This paper offers a political-economy model which provides an intuitive explanation for these observations. In the model, to avoid sanctions-induced challenges, autocratic regimes lower the supply of government services in order to reduce private-sector productivity and hence the resources of potential challengers. This defense strategy only stops working if the sanctions reach a critical intensity. Yet, the critical level might be so high that – even if the regime were ousted and democracy established – imposing sanctions would not be in the interest of the general population.

1 Introduction

Over the course of the 20th century, international economic sanctions have become an increasingly important foreign policy tool. Since the outbreak of World War I, there has been a total of 187 sanctions episodes, about 66 of which started after the collapse of the Soviet empire (Hufbauer et al., 2007; Tables 1A.1 and 1A.2). Economic sanctions usually combine restrictions on international trade and investment and are generally viewed as an instrument to induce specific changes in a target country. In practice, sanctioning states have indicated a variety of goals but the most frequent by far is to promote democratization by pushing autocratic (or even despotic) regimes out of power (46% of the cases in the 1914–2000 period).

So far, however, our knowledge about how economic sanctions might foster *regime change and democratization* is very limited. There is just a general notion that, as Mack and Khan (2000) put it, "the pain inflicted by sanctions on citizens of target states will cause them to pressure their government into making the changes demanded by the sanctioning body." But very little analytical work has actually been devoted to the exact channels through which sanctions are supposed to promote democratization. As a result, our understanding of the factors determining the likelihoods of success and failure is highly incomplete. It is the purpose of the present paper to make some progress in this direction by building a simple political-economy model which reflects the basic features of a typical target country.

Past experiences with the effectiveness of sanctions leave no doubt that a better knowledge about their potential and limitations is required. Moreover, available evidence suggests that sanctions aiming at regime change and democratization have a particularly poor track record: According to Hufbauer et al. (2007), 35% of all the sanctions episodes in the 1914–2000 period can be labeled as at least "partly successful" – which means that the stated goal was at least "partly" achieved and that the sanctions' contribution was at least "substantial". The success rate is much lower, however, when we exclusively consider sanctions imposed to promote regime change and democratization in autocracies. In total, there were 57 such episodes but only 12 of them (or 21%) can be judged as at least "partly successful".[1] In the remaining 79% of the episodes, the imposed measures failed to promote their stated goal. In five cases (or 9%), it appears that they even made a negative contribution.

Recent anecdotal evidence further challenges the simple view on how sanctions should work and highlights the need for a deeper understanding of the underlying mechanisms. As discussed

---

[1] A regime is classified as autocratic if the Polity score (Marshall and Jaggers, 2008) is zero or lower. Note that in 23 cases sanctions to promote regime change were imposed on fairly democratic states with a Polity score above zero. I do not look at these cases since I am interested in the impact of sanctions in autocracies.

in the following section, the pain inflicted on ordinary citizens by the sanctions against Haiti (1991–1994) and Iraq (1990–2003) did not force the governments to give in. Quite the contrary: Substantial evidence suggests that, after the sanctions had been imposed, both governments pursued policies which severely *compounded the economic hardship* caused by the trade and investment embargos. Referring to the situation in Iraq, Mueller and Mueller (1999, p. 49) even note that "the country's political leadership sometimes seems more interested in maximizing the nation's suffering (...) than in relieving it." Similar evidence can be found in other target countries. For instance, in February 2002, Zimbabwe's ruler responded to the threat of sanctions by "promising" to accelerate and broaden the implementation of his disastrous economic policies (*The Economist*; "Zimbabwe on the brink" and "The friendlessness of Robert Mugabe", Feb 19 and Mar 21, 2002). Clearly, such behavior requires explanation.[2]

The political-economy model I am proposing to study the impact of sanctions rests on three natural elements. First, consistent with the focus on regime change and democratization, I consider an autocratic target country, i.e., a country where the government has substantial leeway to implement its preferred policies but also to divert public resources for its own benefit. Second, the state plays an important role in the private sector of the economy: By providing public goods, the government can affect the productivity of private firms and hence the citizens' incomes. Third, challenging the incumbent government in order to promote a transition to democracy imposes an economic cost on the citizens. Interestingly, in such an environment, intensity and effectiveness of sanctions are related in a non-monotonic fashion: As the intensity increases from a low level, matters are initially bound to get worse in the sense that the sanctions fail to promote democratization but just push policies farther away from the democratic ones – which is exactly what we observed in Haiti and Iraq. Only if the coercive measures reach a critical – potentially very high – intensity will economic sanctions become effective in the sense that they start to destabilize the incumbent regime.

The intuition behind this non-monotonicity result is easy to convey: As intended, the imposition of sanctions makes a switch to democracy more rewarding to the ordinary citizens; hence, the sanctions may render the previously reluctant citizenry more inclined to engage in revolutionary activity. Yet, this is exactly what the ruler wants to prevent by making revolutionary activity *more costly* – and a straightforward way to do so is to decrease the supply of public goods: With a lower level of public goods, the citizens' incomes are lower

---

[2]Obviously, regimes may behave like this for propaganda purposes, i.e., to highlight how "inhumane" the sanctions are. However, this cannot be the entire story since the harmful responses are usually not restricted to easily accessible areas but can also be found in the hinterland which are rarely visited by journalists.

and, as a consequence, a given cost associated with an insurgency translates into a bigger loss in terms of instantaneous utility. But as soon as the supply of the public goods hits the lower bound, this defense strategy stops working and a further tightening of the sanctions must destabilize the regime. The evidence I am going to discuss below suggests that the Haitian and Iraqi governments were quite successful in applying such a defense strategy. In particular, it appears that the Iraqi government's strategy to make the middle class struggle was instrumental in preventing serious challenges.

The present analysis also makes a number of comparative-static predictions which might be interesting from a policy perspective, in particular when it comes to assessing the effectiveness of impending sanctions measures. Among other things, the model implies that the minimum intensity to make sanctions work is lower if the state's institutional capacity is higher, i.e., if the state is efficient at taxing economic activity and providing public goods. The intuition is that, if the state's ability to tax economic activity is strong, living under a self-interested dictator carries a higher cost. Thus, the incentives to get rid of the dictator are stronger. Similarly, if the state is more productive in providing public goods, the improvement in public investment associated with a transition to democracy will result in a stronger increase in incomes. Again, this means that the incentives to get rid of the dictator are stronger.

There are few theoretical papers dealing with the impact of sanctions on the political struggle in autocracies. Kaempfer and Lowenberg (1992), for instance, analyze the role of sanctions in helping the domestic opposition to overcome the free rider problem associated with revolutionary activity. More closely related, starting from Wintrobe's (1990) dictatorship model, Kaempfer et al. (2004) focus on the political impact of economic sanctions placed on autocracies in order to alter certain undesirable policies. Finally, Lektzian and Souva (2007) explore why sanctions are less likely to succeed against autocratic targets than against democratic ones. None of these contributions, however, explicitly deals with regime change and democratization nor does any of them address why threatened autocrats may find it optimal to compound the hardship inflicted by the sanctions.[3]

The rest of the paper is organized as follows. Section 2 reviews anecdotal evidence documenting how threatened regimes responded in an extremely harmful manner to the imposition of sanctions. Section 3 introduces the basic model, which is then solved in Section 4. In Section 5, I present an interesting extension of the model which allows for heterogeneity among ordinary citizens. Section 6 discusses policy implications and Section 7 concludes.

---

[3]In an empirical paper, Marinov (2005) looks at whether economic sanctions destabilize country leaders. This does seem to be the case, though to a much smaller extent in autocracies.

2 Anecdotal Evidence from Haiti and Iraq

This section reviews some anecdotal evidence from two recent target countries, Haiti (1991–1994) and Iraq (1990–2003). Besides being well-studied cases, these two sanctions episodes have in common their prime purpose, namely ousting a highly autocratic regime in order to reinstate (Haiti) or establish (Iraq) democracy. In each case, before outlining the regime's policy responses, I will briefly review some facts about the imposed sanctions.

Haiti (1991–1994). In 1990, Haiti saw for the first time fair and democratic presidential elections. The vote produced a clear winner, Jean-Bertrand Aristide, who was sworn into office in February 1991 (see, e.g., Werleigh, 1995, for details). The new president promised to eliminate power and privileges of the old political elites. Yet, this process came to an abrupt halt when a military coup ousted Aristide after just seven months in office. In response, the Organization of American States, including the United Sates, imposed economic sanctions. Initially, these non-mandatory sanctions were targeted at the Haitian government but subsequently became more comprehensive and included severe limitations on most imports and exports from Haiti. Eventually, in June 1993, the United Nations imposed mandatory sanctions, comprising inter alia an oil and arms embargo. However, even with U.N. backing, the sanctions did not achieve their goal; in the end, it was a U.S. military operation that restored democracy in September 1994. The sanctions were lifted in October.

During the time the sanctions were in place, economic activity declined significantly (according to the World Bank's WDI, by about 10% p.a.) and the state almost ceased to function: As noted by Elizabeth Gibbons (1999, p. 31), then UNICEF's representative in Haiti, the new constitutional government "found a thoroughly debilitated, atrophied state structure" when taking over in October 1994. Moreover, Gibbons points out that "the state infrastructure was even more dilapidated than it had been in 1991". It is clear that the sanctions directly challenged the state's ability to perform its functions.[4] However, the rulers' own behavior significantly contributed to this decline. Evidence for the authorities' destructive role comes, inter alia, from the agriculture and infrastructure sectors: Werleigh (1995, pp. 166–7), for instance, points out that the authorities systematically destroyed part of the agricultural infrastructure and prevented technicians from offering their services to the farmers. In addition,

---

[4] For instance, Gibbons (1999, p. 19) argues that the petrol embargo forced schools to operate only three days a week because teachers could not afford transportation to work. Some observers, however, are not convinced that direct effect of the embargo was substantial. They point out that embargo was only weakly enforced, partially lifted from time to time, and excluded the necessities of life (see, e.g., Werleigh, 1995, pp. 164–5).

by arbitrarily expropriating the farmers' returns, they "wiped away any incentive to invest in economic activity." There is also evidence that hundreds of thousands of skilled farmers were forced to abandon their working places, causing a further reduction in agricultural production. Similarly, Gibbons and Garfield (1999, p. 1501) report that the authorities severely hampered the functioning of local water management committees and, as a consequence, compounded the supply problems stemming from the embargo.

Similar examples of harmful government policies can be found in the public health sector. In 1991, after decades under kleptocratic rule, the public health system had already been in a poor state (obviously, the Aristide government could not improve things in just seven months). Yet, even when compared to the dire standards of preceding autocratic regimes, the policies during the sanctions period were extremely poor. In particular, there is evidence that the new rulers deliberately weakened the battered health system even further: Gibbons (1999, p. 49), for instance, reports that there were occasions when the the authorities deprived the entire system of a broad set of vital supplies by impounding medicine, humanitarian fuel, etc.; further examples include the hasty abandonment of vital immunization programs. More generally, Gibbons (1999, p. 50) concludes that the government's (and the military's) policies in the public health sector were extremely harmful, thereby "directly contributing to the violations of the population's basic right to life and health." This assessment is supported by other local observers, among them Farmer et al. (2003, p. 420), who note that the "sharp fall in the quality and coverage of [health] services" was largely due to the striking "absence of commitment to public health on the part of the Haitian army."

In sum, the evidence suggests that government policies during the sanctions period were even poorer than those pursued during the three decades (1957–1986) of personal rule by the Duvalier family. There is no doubt that, by contributing to the collapse of public services and the country's infrastructure, Haiti's military and the various de facto governments in office between 1991 and 1994 considerably deepened the hardship caused by the sanctions.

Iraq (1990–2003). The sanctions against Iraq were imposed by the United Nations Security Council in August 1990 in response to the Iraqi invasion of Kuwait. The sanctions regime was comprehensive but did not include medicines, health equipment or foodstuffs (see, e.g., Reuther, 1995, for details). Moreover, in August 1991, the Security Council authorized the sale of oil in order to pay for the import of humanitarian goods; yet, this opportunity was not taken up by the Iraqi government before 1996. The sanctions had two goals, impairing the country's military potential and promoting regime change and democratization. While the first goal

6

was probably achieved, the second clearly was not. Once again, it took a U.S.-led military operation to oust the incumbent regime. After that, in May 2003, the sanctions were lifted.

When it comes to government policies during the sanctions episode, there are many parallels to the Haitian case. With a comprehensive embargo in place (and the infrastructure damaged by the Gulf War), it was clearly difficult to keep public services going. Yet, many observers emphasize that the government actively contributed to worsening the situation. Reuther (1995, pp. 126, 130), for instance, reports that in large parts of the country the Iraqi regime did not even try to maintain the basic functions of the state or to repair the damaged infrastructure. Instead, the rulers deliberately deprived a substantial fraction of the population of government services and participation in the economy. About 3.1 million people in the three northern governorates were even subjected to a "total internal embargo" which came on top of the international (i.e., "external") sanctions. This "double embargo", jointly imposed by the international community and the Iraqi government, triggered a collapse of economic activity in the north of the country and – as a consequence – led to years of internal conflict between different rival Kurdish factions.[5] Clearly, this ongoing infighting substantially weakened the capacity of the Kurds to challenge the incumbent government in Baghdad.

As in the Haitian case, the public health sector provides further examples of destructive government behavior. For instance, in an article for the *New York Times* ("Were Sanctions Right?", July 27, 2003), David Rieff reports that – although there were more than enough drugs for the political elites – the authorities deprived the population of medical supplies so that ordinary citizens "had been subjected to two sets of sanctions, those of the United Nations and those of Saddam Hussein himself."[6] Such observations are also mirrored in a remark by Hans von Sponeck, then the U.N. coordinator for humanitarian assistance, who – according to the above-mentioned newspaper article – highlighted that "local repression and international sanctions became brothers-in-arms in their quest to punish the Iraqi people for something they had not done." Anecdotal evidence of this type may also constitute the basis of Mueller and Mueller's (1999, p. 49) conclusion that, after the imposition of sanctions, the Iraqi rulers were more interested in maximizing the nation's suffering than in relieving it.

Gibbons (1999, p. 39), finally, argues that the hardship caused by both the sanctions and the government's harmful policies decisively weakened the population's ability to challenge the Iraqi regime. This claim is supported by another report from Iraq in the *New York Times*

---

[5] In addition to strangulating the economy, the regime cut off the region from food aid (Clawson, 1993, pp. 40-1). More generally, it seems that food supplies were diverted away from those who opposed the regime.

[6] According to Garfield (1999), Iraq had invested heavily in health services in the 15 years prior to the embargo and maintained an advanced public health system before the sanctions hit in 1991.

("As Hussein Builds, His People Struggle To Live", January 31, 1998) which cites a diplomat as saying that "if any sector of society outside the military might have formed a political opposition, the Iraqi middle class would have been the only hope." Yet, so the diplomat continues, "it has now been totally destroyed." What the diplomat meant was that most members of the middle class had to assume two or even three jobs to support their families; yet, in spit of these huge efforts, many families were not able to survive without food rations from government. Obviously, this daily struggle to make ends meet, combined with the dependence on the government, left little room for forming or joining an opposition movement.

3 The Model

Agents, preferences, and economic activity. I consider an infinite-horizon economy in discrete time. The society starts out with two different groups of agents, $i \in \{0, 1\}$. Group 0 constitutes the ruling elite and is of an arbitrarily small but positive measure ε. Group 1 consists of ordinary citizens and its size is normalized to unity. In what follows, I completely abstract from within-group differences and assume that all members of a specific group act in cohesion. As a result, each group can be treated as a single actor.

All individuals derive utility from consumption of a non-storable final output good (which also serves as the numéraire). Preferences are given by the intertemporal utility function

$$U_{i,t} = E_t \left\{ \sum_{s=0}^{\infty} \beta^s \ln c_{i,t+s} \right\}, \tag{1}$$

whereas $c_{i,t}$ refers to consumption by the representative member of group $i \in \{0, 1\}$ in period t and $\beta \in (0, 1)$ denotes the discount factor. Note that the instantaneous utility function is logarithmic. This feature will be important for the results as it implies that the marginal utility of consumption becomes arbitrarily large as consumption goes to zero.

The final good is produced by the citizens only. In particular, each citizen has access to a technology which allows to generate a net income (i.e., output minus cost of inputs) of

$$y_t = A_t X_t \tag{2}$$

units of the final good. The first factor in (2), the productivity variable A_t, is taken to reflect the "availability" of crucial foreign input factor. It also serves as the channel through which economic sanctions affect the domestic economy.[7] In particular, I assume that the imposition

---

[7]Note that (2) can be derived from a production function of the form $q_t = ((1-\alpha)^{-(1-\alpha)}\alpha^{-\alpha})(m_t)^{\alpha}(X_t)^{1-\alpha}$, $0 < \alpha < 1$, whereas m is the quantity of the foreign input (whose price, f, depends positively on the sanctions). The optimal choice of m leads to $y_t = AX_t = f^{-\alpha/(1-\alpha)}X_t$, whereas A is a decreasing function of f.

of trade sanctions increases the cost of crucial foreign input factors and hence decreases their use – which is mirrored in a lower net income. The second factor, X_t, refers to the level of the public good provided by the government. It captures that the state plays an important role in promoting private economic activity by, for instance, building and maintaining infrastructure, upholding law and order, or enforcing private contracts.

Figure 1 here

Policies choices and the production of the public good. In each period t, two policy variables have to be determined. First, there has to be a decision on the tax rate on the citizens' incomes. The tax rate is denoted by τ_t and cannot exceed a certain upper bound, $\delta < 1$. Limiting the government's ability to tax is necessary to generate interesting implications. Yet, imposing a maximum tax rate is just a reduced-form way of modeling more realistic limits to taxation which prevent the state from fully appropriating private incomes.[8]

The second policy choice is the investment in public goods. The level of investment is denoted by G_t and the associated cost (in units of the final good) is given by $A_t G_t$. Although not explicitly modeled, an intuitive way of looking at the variable G_t is to assume that it mirrors, for instance, the number of government officials who are in the business of producing public goods. From this perspective, the cost can then be interpreted as the public wage bill which moves in lockstep with private-sector productivity and incomes.

The technology relating public investment, G_t, to the level of the public good is given by

$$X_t = \max \left\{ F(G_t) - \rho_t \chi, 0 \right\}. \tag{3}$$

There are two additive components in (3). The first component is the function F which is increasing and concave in G_t. The function is depicted in Figure 1 and defined according to

$$F(G) = \begin{cases} F^l & : & G^l \le G < G^m \\ F^m & : & G^m \le G < G^h \\ F^h & : & G^h \le G < \infty \end{cases}, \tag{4}$$

with $G^l = 0$ for simplicity. Note that imposing a step function is just for analytical convenience and could be relaxed. The second component in (3), $\rho_t \chi$, reflects that part of the public good will be destroyed if the citizens decide to revolt against the elite (which, as detailed further below, is indicated by $\rho_t = 1$). This is an obvious assumption. Myriad examples suggests

---

[8] In reality, raising taxes is costly, and disproportionally so for high tax rates (as, e.g., modeled in Acemoglu and Robinson, 2001). Thus, the revenue-maximizing rate (which is an interpretation of δ) lies below 100%.

that, when protesters clash with the regime, roads are blocked, law and order collapses, and enforcing contracts turns very difficult. Moreover, it is natural to assume that, as implied by (3), the relative size of the destruction is bigger if the supply of the public good is lower.[9]

As to the relation between the social benefit and the cost of the public good, I assume that

$$F^h - G^h > F^m - G^m > F^l - G^l > 0. \tag{5}$$

Together with (2), (3), and (4), the above condition immediately implies that the surplus-maximizing level of G_t is given by G^h. Yet, G^h is not necessarily the level preferred by the elite. Since it can appropriate at most a fraction δ of the private-sector output, the elite must be interested in the relationship between $\delta A_t F^j$ and $A_t G^j$. In this regard, I assume

$$\delta F^m - G^m > \delta F^j - G^j > 0, \; j \in \{l, h\}, \tag{6}$$

which will ensures that, in absence of an impending revolt, the elite's preferred level of the public good is G^m. Figure 2 illustrates (5) and (6). It shows how productive public investment relates to both private-sector income (Panel a.) and maximum tax revenues (Panel b.).

Figure 2 here

Political regimes and the transition of political power.   There are two political regimes, dictatorship and democracy. In what follows, I denote the political state by $S_t \in \{R, D\}$, whereas D stands for democracy. Under dictatorship, policies are set by the "small" group of elite agents (i.e., by the minority). Moreover, the elite can appropriate and consume any fraction of the tax revenues it likes. Democracy, on the other hand, means that the policies are determined by the "large" group of ordinary citizens (i.e., by the majority).

The economy starts as a dictatorship ($S_0 = R$). However, as long as $S_t = R$, the elite's power is continuously threatened. More specifically, in each period t, the citizens may decide to oust the ruling elite. The citizens' decision in this respect is denoted by $\rho_t \in \{0, 1\}$, with 1 indicating a "revolt". If the citizens indeed choose $\rho_t = 1$, democracy will be *irreversibly* established in the following period (i.e., $S_{t+1} = S_{t+2} = \cdots = D$).[10] Moreover, the elite is no longer part of the game from $t + 1$ onwards but each member receives a continuation value of $W(D) = \ln(\omega/\varepsilon)(1 - \beta)^{-1}$. It is most natural to think of ω as the recurrent payoff from an asset which – independently of the political state – is in possession of the elite.

---

[9]If, e.g., the traffic infrastructure is underdeveloped, one blocked road may be sufficient to cause gridlock. Similarly, an understaffed judiciary is highly vulnerable to absenteeism (associated with political turmoil).

[10]Imposing D to be an absorptive state is not crucial for the results, however. In a simple extension of the model – which includes a return to R with probability $q > 0$ in each period – the results are very similar.

It is clear, though, that the political state is unchanged if the citizens abstain from a revolt ($\rho_t = 0$). Abstaining may indeed be an option because, as discussed above, staging a revolt means that part of the public good is destroyed – which decreases individual incomes.

Finally, when ousted, the elite loses part of the appropriated resources, with the loss increasing in public investment (because escaping a better funded law-enforcement apparatus is more costly, for instance). To keep matters simple, I assume that the levels G^m and G^h lead to a full loss whereas G^l implies a partial loss. Though reasonable, these assumptions are just for convenience and irrelevant to the model's qualitative implications.[11]

Economic sanctions. The sole aim of the sanctioning body – which is called the sender state for simplicity – is to induce regime change and democratization (i.e., a transition from state R to D). To pursue this goal, the sender state may impose sanctions which deplete the productivity of the domestic economy. More specifically, the sender state can push the productivity parameter A_t below its "natural" level which is normalized to 1. In reality, such a negative impact on productivity is usually achieved by imposing restrictions on international trade, especially the denial of critical imports (see, e.g., Hufbauer et al., 2007, pp. 44-5). It is obvious, though, that the import of critical input factors can never be perfectly prohibited (for instance, due to the possibility of smuggling or because some trading partners refuse to impose sanctions). In practice, as already alluded to above, the denial of critical imports means that these goods become more expensive and hence are used in lower quantities.

As for the sanctions strategy, I assume that the only variable the sender state can condition on is the political state, S_t. This is reasonable because it is difficult for a foreign authority to exactly observe the targeted regime's policy choices or the responses by the opposition (which are the model's remaining endogenous variables). Moreover, even if this were possible, the foreign decision makers (e.g., the U.S. Congress or the U.N. Security Council) would hardly have the capacity to review and adjust sanctions decisions after every single policy decision in the target country. Thus, it seems natural to focus on strategies of the form

$$A_t = \begin{cases} 1 & : \quad S_t = D \\ \underline{A} & : \quad S_t = R \end{cases},$$

whereas $\underline{A} < 1$. Moreover, for further use below, it is convenient to introduce $\Delta \equiv 1/\underline{A} \in [1, \infty)$ as a measure for the *intensity* of the sanctions regime imposed.

---

[11]In particular, they ensure that the elite prefers the lowest level of investment, G^l, when anticipating a revolt – which, in turn, implies that there is always a pure strategy Markov Perfect Equilibrium. Without these assumptions, there would be parameter constellations for which only a mixed strategy equilibrium exists.

Equilibrium concept and time-line. The focus is on the (pure strategy) Markov Perfect Equilibrium (MPE), where strategies depend only on the payoff-relevant states and prior actions within the same period. In the present setup, the only state variable is S_t.

Suppose first that $S_t = R$. Then, the timing of events is as follows: *First*, the elite determines the policy vector, $\Pi_t = (\tau_t, G_t)$. *Second*, the citizens decide whether or not to oust the regime. *Third*, all decisions are implemented, the payoffs materialize, and the period ends. If the elite has not been challenged ($\rho_t = 0$), the political state remains unchanged (i.e., $S_{t+1} = R$). Otherwise, if $\rho_t = 1$, the elite is ousted at the end of the period and the country turns into a stable democracy (i.e., $S_{t+1} = S_{t+2} = \cdots = D$).

Suppose now that $S_t = D$. Then, the citizens determine Π_t, these policies are immediately implemented, the payoffs materialize, and the period ends (whereas $S_{t+1} = D$).

4   Analysis

This section derives the politico-economic equilibrium and explores the impact of sanctions. I first assume that the state's repressive capacity is "intermediate" in the sense that

$$\phi F^l < \chi < F^l, \tag{7}$$

whereas $\phi \equiv ((F^m/F^l)^{1/(1-\beta)} - F^m/F^l)/((F^m/F^l)^{1/(1-\beta)} - 1)$. This is the constellation under which the model generates the richest set of implications. After that, the cases of "high" (i.e., $F^l \leq \chi$) and "low" (i.e., $\phi F^l \geq \chi$) repressive capacities are easily discussed.

4.1   Intermediate Repressive Capacity

4.1.1   The Equilibrium under Democracy

It is convenient to derive the equilibrium under democracy first. To do so, note that the government budget constraint under democracy is simply given by $\tau_t F(G_t) \geq G_t$ because the elite is no longer part of the game and hence no longer consumes or faces challenges. Moreover, the constraint must hold with equality since imposing taxes higher than necessary to finance public investment is clearly suboptimal. Thus, given the size of the public investment, we have $\tau_t = F(G_t)/G_t$ so that consumption by the representative citizen is given by

$$c_{1,t} = F(G_t) - G_t.$$

Clearly, because of (4) and (5), maximizing the above expression requires $G_t = G^h$. As a result, all citizens agree on the policy vector $\Pi_t = ((F^h/G^h, G^h)$. Moreover, since switching

back to dictatorship is impossible, identical policies will be implemented in all future periods $t + 1$, $t + 2$, $\cdots$. Thus, once the political state has switched to D, the uniform level of lifetime utility incurred by each citizen is given by

$$V(D) = \frac{\ln(F^h - G^h)}{1 - \beta}. \tag{8}$$

4.1.2 The Equilibrium under Dictatorship

Some basic properties of the equilibrium. I now describe the MPE under dictatorship and start by deriving some basic properties. To do so, remember that the elite has to observe a number of restrictions when deciding on $\Pi_t = (\tau_t, G_t)$. The first one is the upper bound on the tax rate, $\tau_t \leq \delta$. The second one is given by the government budget constraint,

$$\varepsilon c_{0,t} \leq \omega + \tau_t A_t X_t - A_t G_t, \tag{9}$$

which states that total consumption by the elite must not exceed the current budget surplus (public revenues minus public investment). Moreover, besides these formal constraints, the elite will take into account the implications of its choices for the citizens' decision on ρ_t.

When deciding on ρ_t, the latter compare the associated costs and benefits. Obviously, the representative citizen is not in favor of a revolt if the cost of ousting the elite exceeds the benefit. In formal terms, given $\Pi_t = (\tau_t, G_t)$, the representative citizen prefers not to revolt if

$$\ln\left((1 - \tau_t)\underline{A}F(G_t)\right) + \beta V(R) > \ln((1 - \tau_t)\underline{A}(F(G_t) - \chi) + \beta V(D),$$

whereas $V(R)$ stands for the citizens' value function under dictatorship and $V(D)$ is given by (8). By rearranging terms, the above condition simplifies to

$$\ln\left(\frac{F(G_t)}{F(G_t) - \chi}\right) > \beta\left(V(D) - V(R)\right). \tag{10}$$

Condition (10) is a key equation in the present analysis. The left-hand side reflects the cost of revolting against the elite in terms of a drop in instantaneous utility: In the case of a revolt, the current income and hence the current level of consumption is reduced. The right-hand side gives the benefit in terms of an increase in the continuation value: Democracy (i.e., the state under which the citizens' utility is maximized) will be permanently established from period $t + 1$ onwards. Thus, the condition states that the representative citizen does not support a revolt if the *short-run* utility cost exceeds the *long-run* benefit.

The following lemma establishes two implications which – independently of the sanctions regime and the exact parameter constellation – can be derived from (9) and (10):

13

Lemma 1 *Suppose that the political state is dictatorship ($S_t = R$). Then, in any Markov Perfect Equilibrium, the ruling elite (i) sets the maximum tax rate (i.e., $\tau_t = \delta$); (ii) does not choose levels of G other than G^m or G^l.*

Proof. See Appendix. ■

The intuition behind the first claim is straightforward: The current tax rate, τ_t, does not enter (10) and hence cannot influence the citizens' current decisions on ρ_t. As a result, whether or not the elite faces a revolt is independent of current taxes – which means that it simply chooses the rate which, other things equal, maximizes the current budget surplus (i.e., the right-hand side of 9). As for the investment in public goods, it is obvious that levels other than G^l, G^m, and G^h cannot be optimal. But why can G^h be ruled out? Intuitively, opting for G^h cannot occur in an MPE because a (one-time) deviation must increase the elite's utility: For instance, if $\rho_t = 0$ is anticipated, choosing G^m over G^h increases the current budget surplus (due to 6) and hence consumption by the elite. But doing so also raises the cost of revolt which is given on the left-hand side of (10). Thus, while strictly improving the elite's current level of consumption, such a deviation cannot trigger a revolt and must therefore be optimal.

Equilibrium 1 (Stable dictatorship with poor economic policies). I now establish the different types of equilibria and discuss how they relate to the intensity of the sanctions regime. The first constellation I look at is $\Delta < \Lambda_1$ ("low" sanctions' intensity), whereas

$$\Lambda_1 = \left(\frac{F^m}{F^m - \chi} \right)^{(1-\beta)/\beta} \frac{(1-\delta)F^m}{F^h - G^h}. \tag{11}$$

Under these circumstances, there exists a MPE where $\Pi_t = (\delta, G^m)$ and $\rho_t = 0$ for all t.

To see this, it is convenient to go backwards through a given period t. So suppose that we are at the point where the citizens have to decide whether or not to oust the elite. Clearly, the citizens' best response to $\Pi_t = (\delta, G^m)$ is to avoid a revolt if and only if (10) is satisfied. To check the validity of this condition, note that – given that the assumed equilibrium policies are implemented in all future periods $t+1, t+2, \cdots$ – we have

$$V(R) = \frac{\ln \left((1-\delta)\underline{A}F^m \right)}{1 - \beta}. \tag{12}$$

Then, taking (8) and (12) into account, the decisive condition for $\rho_t = 0$ turns into

$$\ln \left(\frac{F^m}{F^m - \chi} \right) > \frac{\beta}{1 - \beta} \ln \left(\Delta \frac{F^h - G^h}{(1-\delta)F^m} \right),$$

which is equivalent to $\Delta < \Lambda_1$. Thus, given that the ruling elite chooses $\Pi_t = (\delta, G^m)$ in the first step, the citizens will actually abstain from a revolt ($\rho_t = 0$) later on.

Moving one step backwards, it remains to analyze the elite's policy choice. Given that (due to 6) the vector $\Pi_t = (\delta, G^m)$ maximizes the elite's current consumption (9), and since this choice is followed by $\rho_t = 0$, it has no incentive to deviate from this policy combination either. Thus, if (and only if) we have $\Delta < \Lambda_1$, this type-1 equilibrium does exist.

Equilibrium 2 (Stable dictatorship with disastrous economic policies). The second constellation is $\Lambda_1 \leq \Delta < \Lambda_2$ ("intermediate" sanctions' intensity), whereas

$$\Lambda_2 = \left(\frac{F^l}{F^l - \chi} \right)^{(1-\beta)/\beta} \frac{(1-\delta)F^l}{F^h - G^h} \tag{13}$$

and $\Lambda_1 < \Lambda_2$ due to the lower bound on χ imposed in (7). Under these circumstances, there exists a MPE where $\Pi_t = (\delta, G^l)$ and $\rho_t = 0$ for all t.

This equilibrium can again be established by backward induction. An approach completely similar to the one above shows that – given that the assumed equilibrium policies are implemented in all future periods $t+1, t+2, \cdots$ – we will have $\rho_t = 0$ in response to $\Pi_t = (\delta, G^l)$ if and only if $\Delta < \Lambda_2$. Moving one step backwards, we further need to check whether the elite is actually best off by opting for G^l in the first step of period t. This is the case if a deviation to $G_t = G^m$ would trigger a revolt against the elite. By looking at the citizens' decision problem once more (condition 10), one can establish that they would be in favor of a revolt if

$$\ln \left(\frac{F^m}{F^m - \chi} \right) \leq \frac{\beta}{1 - \beta} \ln \left(\Delta \frac{F^h - G^h}{(1-\delta)F^l} \right),$$

which is equivalent to $(F^l/F^m)\Lambda_1 \leq \Delta$. But this latter inequality must be satisfied since $(F^l/F^m) < 1$ and the focus is on the case $\Lambda_1 \leq \Delta < \Lambda_2$ ("intermediate" sanctions' intensity). Thus, if Δ lies in the above-mentioned range, the type-2 equilibrium does exist.[12]

Equilibrium 3 (Challenged dictatorship). Finally, in the case of $\Lambda_2 \leq \Delta$ ("high" sanctions' intensity), there is a MPE where $\Pi_t = (\delta, G^l)$ and $\rho_t = 1$ (so that $S_{t+1} = D$).

To see this, let me again focus first on the citizens' decisions on whether or not to revolt against the elite. Clearly, the citizens' best response to $\Pi_t = (\delta, G^l)$ is to oust the elite if and only if (10) is violated. To check this, note that – given that the assumed equilibrium policies are implemented in all future periods $t+1, t+2, \cdots$ – we have

$$V(R) = \ln \left((1-\delta)\underline{A}(F^l - \chi) \right) + \beta V(D). \tag{14}$$

Then, taking (8) and (14) into account, the decisive condition for $\rho_t = 1$ becomes

$$\ln \left(\frac{F^l}{F^l - \chi} \right) \leq \beta \ln \left(\Delta \frac{F^h - G^h}{(1-\delta)(F^l - \chi)} \right).$$

---

[12] More precisely, the type-2 equilibrium does exist if (and only if) we have $(F^l/F^m)\Lambda_1 \leq \Delta < \Lambda_2$.

It turns out that this condition is satisfied if $\Lambda_2 \le \Delta$ – which was exactly imposed above.

Moving one step backwards, it is clear that – given that the assumed equilibrium policies are implemented in all future periods $t+1, t+2, \cdots$ – the elite has no incentives to deviate from $\Pi_t = (\delta, G^l)$: Opting for either G^m or G^h would just decrease the utility cost of a challenge and hence could not make the citizens abstain from a revolt. Thus, the ruling elite inevitably faces a revolt later on in the period – which means that G^l is the preferred option. Thus, if (and only if) we have $\Lambda_2 \le \Delta$, the type-3 equilibrium does exist.

Further results and summary. Two additional points are worth noting. First, there are no further equilibria: Lemma 1 rules out all policy vectors other than (δ, G^l) or (δ, G^m); moreover, a MPE where $\Pi_t = (\delta, G^m)$ and $\rho_t = 1$ (as long as $S_t = R$) cannot exist since – in anticipation of a revolt – the elite prefers G^l. The second point is that the use of sanctions cannot improve the citizens' welfare: In a given equilibrium, $V(R)$ decreases monotonically as Δ goes up; moreover, $V(R)$ drops discontinuously as soon as Δ reaches Λ_1 (since public investment falls) and does not improve when Δ reaches Λ_2 (since at that point the citizens are indifferent between revolting or not). Note, however, that Section 5 revisits the welfare issue in a slightly extended version of the model which allows for heterogeneity among the citizens.

Figure 3 here

Figure 3 above gives a graphical illustration of how the intensity of the sanctions relates to the three different types of equilibria. A complete overview of the model's implications is presented in the following proposition (proof in the text):

Proposition 1 *Suppose that the political state is dictatorship ($S_t = R$). Moreover, assume $\phi F^l \le \chi \le F^l$ and $\Lambda_1 > 1$. Then, depending on the intensity of economic sanctions, the following three types of Markov Perfect Equilibria can exist:*

- *Equilibrium 1: If $\Delta < \Lambda_1$ ("low" intensity), there is an MPE where $\Pi_t = (\delta, G^m)$ and $\rho_t = 0$ for all t (and there is also Equilibrium 2 if $(F^l/F^m)\Lambda_1 \le \Delta$).*

- *Equilibrium 2: If $\Lambda_1 \le \Delta < \Lambda_2$ ("intermediate" intensity), there is a unique MPE where $\Pi_t = (\delta, G^l)$ and $\rho_t = 0$ for all t.*

- *Equilibrium 3: If $\Lambda_2 \le \Delta$ ("high" sanctions), there is a unique MPE where $\Pi_t = (\delta, G^l)$ and $\rho_t = 1$ (so that $S_{t+1} = D$).*

Moreover, independently of the parameter constellation or the intensity applied, the use of sanctions always reduces the welfare of the ordinary citizens in the target country.

Proposition 1 presumes $\Lambda_1 > 1$. It is clear, though, that this condition does not necessarily hold. On the one hand, it is possible that the actual parameter constellation gives rise to $\Lambda_1 \leq 1 < \Lambda_2$. In this case, even in the absence of any sanctions, the economy would be in equilibrium 2 – and a rising sanctions intensity would just take it from there to equilibrium 3. On the other hand, if $\Lambda_1 < \Lambda_2 \leq 1$, the elite would be ousted even if there were no sanctions imposed. It is the purpose of the following subsection to discuss how the different parameters affect the $\Lambda-$thresholds and hence the sequence of equilibria.

4.1.3   Discussion of Results

Sanctions' intensity and effectiveness. Proposition 1's central implication is that – as illustrated in Figure 4 – intensity and effectiveness of sanctions may be related in a non-monotonic way. A low sanctions' intensity may be ineffective in the sense that it fails to promote democratization. Yet, a low intensity does not cause additional damage beyond the direct negative impact of the sanctions. An intermediate intensity, though, may not only be ineffective but detrimental in the sense that it causes additional damage by permanently pushing policies farther away from those preferred by the population (i.e., from "poor" to "disastrous"). Finally, imposing a high intensity may be effective in the sense that it does promote democratization and a higher supply of the public good in all future periods.

Figure 4 here

This non-monotonicity result is a direct consequence of the elite's defense strategy. The ruling elite does not face a revolt as long as the citizens feel that the (utility) cost of a challenge is "high" relative to the expected benefit. As the sanctions regime becomes more intense, though, the expected benefit from a switch to democracy rises. Thus, at some point, the elite has to increase the cost side in order to stay unopposed. This need to intensify the pain explains exactly why domestic policies may worsen in response to tighter sanctions: By providing a lower level of the public good, the elite decreases the citizens' current income and hence increases the marginal product of consumption; as a result, the loss in income induced by a revolt translates into a bigger cost in terms of lower instantaneous utility from consumption. Yet, as soon as the supply of the public good hits the lower bound, this strategy of deliberately worsening economic policies has reached its limits. Then, the pain can no longer be increased and a further tightening of sanctions will inevitably destabilize the regime.

It is finally interesting to note that – given the assumed parameter constellation holds – the elite's defense strategy works despite the fact that it leads to a lower supply of the public good

in all future periods and hence to a bigger expected benefit from revolting against the regime. Obviously, then, it must be the case that the increase in the cost of a revolt outweighs the rise in the benefit. The main reason is the shape of the instantaneous utility function: Since instantaneous utility is concave in consumption, reducing the supply of the public good must increase the utility cost of a revolt by a larger amount than it widens the difference between the levels of the instantaneous utility under, respectively, dictatorship and democracy. A further reason is that the cost of a revolt materializes immediately while the benefits set in only in the future and hence are discounted. Therefore, depending on the size of the discount factor, even a possibly steep increase in the gap between the instantaneous utilities associated with R and D may lead to a relatively mild rise in the future benefit from a revolt.

Comparative-static results. We will see that the minimum intensity to make sanctions work, Λ_2, depends in a clear-cut way on the state's *repressive* and *institutional* capacities. The state is said to have a strong repressive capacity if it is good at fighting internal challenges so that the cost of successfully ousting the regime is high. In the present setup, this cost is mirrored by χ. The parameter χ gives the size of the damage done to the public good and thus can be viewed as a measure for the intensity of the struggle. A strong institutional capacity, on the other hand, means that the state is efficient at taxing economic activity and providing the public good (see, e.g., Besley and Persson, 2009). These aspects are captured by the parameter δ and the ratio $(F^h - G^h)/F^l$ which reflect, respectively, the maximum tax rate and the productivity of the public goods technology.

Looking at (13), we see that Λ_2 is *increasing* in the repressive capacity of the state. Clearly, the higher the cost of a revolt, the higher the sanctions-induced pain required to make the citizens revolt. The critical threshold is *decreasing,* though, in the state's institutional capacity (as proxied by δ and $(F^h - G^h)/F^l$). With a stronger tax bureaucracy, dictatorship is more painful to the citizens since the ruling elite is able to appropriate a larger fraction of their incomes. Thus, under these circumstances, a switch to democracy is more rewarding. Similarly, if the state is more productive in providing public goods, the rise in public investment associated with a transition to democracy will result in a stronger increase in incomes. Thus, again, we get the result that a switch to democracy is more rewarding.

4.2 High or Low Repressive Capacity

This section closes with a brief discussion of what changes if (7) is violated. The overall finding is that the set of possible equilibria is more limited. This is most obvious in the case of a

high repressive capacity. Clearly, if $F^l \leq \chi$, equilibrium 3 cannot exist because a revolt would destroy the entire public good (equations 3 and 4) and hence reduce the citizens' incomes to zero (equation 2). As a result, instantaneous utility would go to minus infinity. In formal terms, if χ approaches F^l, the critical threshold Λ_2 goes to infinity which means that no (finite) sanctions' intensity is sufficient to induce a revolt. Thus, under these circumstances, all economic sanctions can achieve is to push the economy from equilibrium 1 to 2.

On the other hand, in the case of a low repressive capacity (i.e., if $\chi \leq \phi F^l$), reducing public investment can no longer prevent a revolt once the sanctions' intensity reaches Λ_1. It is still true that a lower supply of the public good increases the instantaneous utility cost of a revolt. However, under these circumstances, this cost increase is insufficient to outweigh the increase in the payoff associated with a switch to democracy. In formal terms, $\chi \leq \phi F^l$ implies $\Lambda_2 \leq \Lambda_1$. As a result, as can be checked along the lines demonstrated above, the economy will no longer jump to equilibrium 2 but will go directly to equilibrium 3 as soon as the sanctions' intensity reaches Λ_1. So imposing economic sanctions does not carry the risk of accidentally pushing the economy into its least satisfactory state.

5   Heterogeneity and Welfare

The baseline version of the model leaves no room for economic sanctions to improve the citizens' payoffs. Yet, this is different in a slightly extended version of the model which allows for heterogeneity among ordinary citizens. In such a modified setup, as this section will show, sanctions may not only be effective but actually raise the average citizen's payoff.

5.1   The Modified Setup

Deviations from the baseline model.   Suppose that the cost of a revolt is no longer equally spread among the citizens. Instead, assume that the magnitude of the disruptive effect varies across different population groups. This is a reasonable assumption since, to give an example, the regime's capacity to fight back might differ across regions. Specifically, suppose that the citizens can be divided into two equal-sized groups so that now $i \in \{0, 1, 2\}$, with 0 still referring to the elite. Without loss of generality, group 1 is assumed to consist of citizens who face an above-average degree of destruction, $\chi_1 \equiv \chi + \gamma$, while group 2 consists of citizens who face below-average destruction, $\chi_2 \equiv \chi - \gamma$. Suppose further that the two groups simultaneously decide on whether or not to revolt against the elite. Each group's decision in this respect is denoted by $\rho_{i,t} \in \{0, 1\}$, with 1 indicating a revolt. However, a revolt is only successful if

supported by both groups ($\rho_{1,t} = \rho_{2,t} = 1$); a revolt initiated by just one group is insufficient to oust the regime. Finally, for the rest of this section, I assume that the state's repressive capacity is intermediate so that condition (7) holds (with χ_i, $i \in \{1, 2\}$, replacing χ).

Equilibria. It can easily be checked along the lines demonstrated in Subsection 4.1 that these mild modifications neither change the properties nor the sequence of the 3 different equilibria summarized in Proposition 1. There are only two slight formal adjustments. First, we now have $\rho_{1,t} = \rho_{2,t} = 0$ (instead of just $\rho_t = 0$) in equilibria 1 and 2 and $\rho_{1,t} = \rho_{2,t} = 1$ (instead of just $\rho_t = 1$) in equilibrium 3. Secondly, χ has to be replaced by χ_1 in the definitions of Λ_1 and Λ_2 (equations 11 and 13). As a result, for a given average cost χ, the minimum intensity required to make the sanctions work is higher if there is heterogeneity.[13] This is intuitive since a successful revolt necessarily requires the support of the high-cost group.

5.2 Welfare Implications

The welfare implications are now readily assessed. As is the case in the baseline setup, a positive effect on the citizens' payoffs can be ruled out if the type-3 equilibrium prevails even in the absence of any sanctions ($\Lambda_2 < 1$). Under these circumstances, the citizens would revolt anyway and get rid of the regime after just one period (so that $S_1 = D$). However, unlike in the baseline version, the average impact on individual payoffs may be positive if the regime is unchallenged in the absence of sanctions:

Proposition 2 *Suppose that $\Lambda_2 \geq 1$. Moreover, note that F^j takes on either the value F^l or F^m depending on the equilibrium that prevails if $\Delta = 1$. Then, we have*

$$E\left\{V_i(R, \Lambda_2) - V_i(R, 1)\right\} = \ln\left(\left[\frac{F^l/F^j}{\Lambda_2}\right]^{1/(1-\beta)}\left[\frac{F^l - \chi_2}{F^l - \chi_1}\right]^{1/2}\right), \quad i \in \{1, 2\}, \qquad (15)$$

whereas $V_i(R, \Delta)$ is the lifetime utility of a member of group i (given state R and intensity Δ).

Proof. See Appendix. ∎

Equation (15) represents the average impact on individual payoffs if the sanctions' intensity increases from 1 to Λ_2, i.e., to the critical level which is exactly sufficient to induce the type-3 equilibrium. Since $\Lambda_2 \geq 1$ and $F^l/F^j \leq 1$, the equation immediately reveals that the average impact can only be positive if the cost of a revolt differs across the two groups (i.e., if $\gamma > 0$ so

---

[13]Moreover, note that there exists an additional MPE where, independently of Δ, the ruler chooses $\Pi_t = (\delta, G^m)$ and, irrespective of the ruler's choice, each of the two groups follows the strategy to opt for $\rho_{i,t} = 0$. Yet, for the rest of this section, I assume that equilibrium 3 arises as soon as Δ reaches Λ_2.

that $\chi_1 > \chi_2$). The intuition is as follows: A revolt against the ruling elite requires the support of both groups. But this means that group 1 (which faces a higher cost) may find it optimal to "dictate" a non-challenge equilibrium although group 2 might strictly prefer the type-3 equilibrium. In such a situation, group 1 does not take into account that group 2 would benefit from a revolt. Thus, from the perspective of the average citizen, the sanctions-free equilibrium may show an inefficiently low level of "revolutionary activity". As a result, a sanctions-induced switch to equilibrium 3 can improve matters. Expression (15) further reveals that the average effect is more likely to be positive if the critical intensity Λ_2 is low (because, for instance, a switch to D leads to a strong improvement in the supply of public goods).

Figure 5 here

Figure 5 gives a graphical illustration of how average individual payoff relates to the intensity of the sanctions imposed. There are two examples, one with $\Lambda_1 > 1$ (Panel *a.*) and the other with $\Lambda_1 < 1 < \Lambda_2$ (Panel *b.*). In both cases, as long as $\Delta < \Lambda_2$, the average payoff decreases as the sanctions become more intense (and even jumps down at Λ_1 in situation *a.*). At Λ_2, it jumps up but turns again into a decreasing function of Δ beyond Λ_2. The spike at Λ_2 occurs because – at this point – group 2 strictly prefers the third over the second equilibrium (while group 1 is indifferent). Finally, whether or not $E\left\{V_i(R, \Lambda_2) - V_i(R, 1)\right\} > 0$ depends on the exact parameter constellations (see Proposition 2). Yet, this inequality is more likely to hold if the parameter constellation implies a "low" Λ_2 (as shown in panel *b.*).

6   Policy Perspective

The above analysis may offer a new perspective on what happened in Haiti and Iraq. Through the lens of this model, it appears that the imposed sanctions were too weak to have a destabilizing effect but sufficient to take the two countries from a type-1 equilibrium to a much worse equilibrium of the second type. Put differently, the intensities were insufficient to fend off the "defense strategy" highlighted in the previous section: By scaling back public services to an extent which exceeded unavoidable adjustments, the two regimes managed to increase the cost of revolting in a way that countered the heightened revolutionary zeal in the populations. As a result, the regime went unchallenged but the citizens were punished twice, through the direct effect of the sanctions and by the regimes' calculated policy responses.

Obviously, from a policy perspective, the model's main implication is that the intensity of the coercive measures plays a central role. There are two different aspects. First, the model

implies that, in order to promote regime change and democratization, the sanctions must be sufficiently severe. If the critical level is missed, even only narrowly, there is nothing to show in return for the hardship inflicted on the population by both the sanctions and the regime's response. Thus, clearly, such sanctions should be strictly avoided if the critical intensity cannot be achieved because, e.g., the available coercive measures are too weak; because they are backed by too few countries; because they cannot be effectively enforced. Moreover, closely related to this finding, the model highlights that an intermediate sanctions' intensity may be the worst thing in terms of welfare. Thus, paradoxically, the present analysis highlights the possibility that diluting sanctions to spare the general population may achieve the exact opposite. The second aspect is that the critical intensity may be so high that, even if the sanctions successfully induce a transition to democracy, everyone's welfare drops in the moment the measures are imposed. So, in spite of bringing democracy, economic sanctions may actually go against the interests of the general population.

The present analysis also points to a number of hitherto less observed factors which might influence the ability of economic sanctions to promote regime change and democratization. In particular, such sanctions are more likely to work (and more likely to be in the interest of the average citizen) in economies which can be considered to be more advanced. In the present setting, the term "more advanced" has two different meanings. One the one hand, it means that the state's institutional capacity is relatively high. If the state is efficient at supply public goods but the self-interested regime – for defense purposes – keeps public investment low, a switch to democracy brings big increases in utility. As a result, mild sanctions might be sufficient to induce democratization. On the other hand, more advanced means that the economy has more interaction with the outside world (e.g., in the context of Footnote 3, a higher α), giving the latter more leeway to impose measures that are sufficiently intense.

7 Conclusions

Economic sanctions aiming at regime change and democratization are frequently imposed, sometimes very harmful, but hardly ever successful. Given these facts, there is surprisingly little research on how such sanctions work – and therefore also little knowledge about why they frequently fail. The present paper starts to fill this gap by analyzing the impact of sanctions in a parsimonious model of nondemocratic politics. The main finding is that targeted regimes may resort to a simple yet powerful defense strategy: By aggressively lowering the supply of public services, they can make the citizens poorer and thus increase the strain associated with

the disruptive effects of a revolt. In the light of the present analysis, it thus appears that the regimes in Haiti and Iraq deliberately amplified the sanctions' negative impact in order to stay in power. This amplification strategy is less likely to work, though, if the sanctions' intensity is high – and herein lies the problem. In practice, agreeing on crippling sanctions is very difficult since such sanctions mean significant hardship for the population and also touch economic interests in the sender states. Thus, crippling measures are likely to be watered down – which may give rise to a most unsatisfactory situation: The sanctions cannot attain their purpose but the citizens are punished twice, not only through the direct effect but also via the regimes' harmful policy responses. Moreover, the intensity required to promote democratization might be so high that – even if it could be reached – everyone in the target country would be worse off. Although the implications of the present theory are consistent with substantial anecdotal evidence, there might be many other factors behind the failure of past sanctions episodes. Thus, more research – both theoretical and empirical – has to be done in order to get a more comprehensive view on the potential and limitations of sanctions.

References

[1] Acemoglu, Daron and James Robinson (2001); "A Theory of Political Transitions", *American Economic Review*, 91(4), 938-963.

[2] Clawson, Patrick (2003); *How has Saddam Hussein Survived? Economic Sanction, 1990 – 1993*. Washington, DC: National Defense University.

[3] Farmer, Paul, Mary Smith, and Patrice Nevil (2003); "Unjust Embargo of Aid for Haiti", *The Lancet*, 361, 420-423.

[4] Garfield, Richard (1999); *The Impact of Economic Sanctions on Health and Well-being*. London: Overseas Development Institute.

[5] Gibbons, Elizabeth (1999); *Sanctions in Haiti: Human Rights and Democracy under Assault*. Westport, CT: Praeger Publishers.

[6] Gibbons, Elizabeth and Richard Garfield (1999); "The Impact of Economic Sanctions on Health and Human Rights in Haiti, 1991-1994," *American Journal of Public Health*, 89(10), 1499-1504.

[7] Hufbauer, Gary Clyde, Jeffrey Schott, Kimberly Ann Elliott, and Barbara Oegg (2007); *Economic Sanctions Reconsidered* (3rd ed.). Washington, DC: Peterson Institute for International Economics.

[8] Kaempfer, William and Anton Lowenberg (1992); "Using Threshold Models to Explain International Relations", *Public Choice*, 73(4), 419-443.

[9] Kaempfer, William, Anton Lowenberg, and William Mertens (2004); "International Economic Sanctions Against a Dictator", *Economics and Politics*, 16(1), 29-51.

[10] Lektzian, David and Mark Souva (2007); "An Institutional Theory of Sanctions Onset and Success", *Journal of Conflict Resolution*, 51(6), 848-871.

[11] Mack, Andrew and Asif Khan (2000); "The Efficacy of U.N. Sanctions", *Security Dialogue*, 31(3), 279-292.

[12] Marinov, Nikolay (2005); "Do Economic Sanctions Destabilize Country Leaders?", *American Journal of Political Science*, 49(3), 564-576.

[13] Marshall, Monty and Keith Jaggers (2008); *Polity IV Project: Political Regime Characteristics and Transitions, 1800-2008*. Electronic resource.

[14] Mueller, John and Karl Mueller (1999); "Sanctions of Mass Destruction", *Foreign Affairs*, 78(3), 43-53.

[15] Reuther, David (1995); "UN Sanctions Against Iraq", in David Cortright and George Lopez, eds., *Economic Sanctions: Panacea or Peacebuilding in a Post-Cold War Word?* Boulder, CO: Westview Press.

[16] Werleigh, Claudette Antoine (1995); "The Use of Sanctions in Haiti: Assessing the Economic Realities", in David Cortright and George Lopez, eds., *Economic Sanctions: Panacea or Peacebuilding in a Post-Cold War Word?* Boulder, CO: Westview Press.

[17] Wintrobe, Ronald (1990); "The Tinpot and the Totalitarian: An Economic Theory of Dictatorship", *American Political Science Review*, 84(3), 849-872.

Appendix

Proof of Lemma 1. Claim (i): Assume first that – contrary to what is stated in the lemma – there exists an MPE where $\tau_t < \delta$ and $\rho_t = 0$ for all t. This assumption can easily be led to a contradiction by looking at the citizens' decision on whether or not to challenge the elite in a certain period t: Given that the assumed equilibrium strategies are applied in all future periods $t+1$, $t+2$, $\cdots$, it must be the case that – when deciding on ρ_t – the citizens are better off by not revolting against the elite. Put differently, it must be the case that condition (10) holds, with the equilibrium value of G on the LHS. Yet, condition (10) does not depend on the current tax rate, τ_t. Thus, even if the elite deviated and chose $\tau_t = \delta$ in the first step of the stage game (instead of setting τ_t equal to the "equilibrium" value), the best response would still be to abstain from challenging the elite. As a result, the elite must find it optimal to choose $\tau_t = \delta$ (i.e., to implement the revenue-maximizing tax rate) – which contradicts the initial assumption. Note that assuming the existence of an MPE where – as long as $S_t = R$ – we have $\tau_t < \delta$ and $\rho_t = 1$ can be led to a contradiction in a similar way.

Claim (ii): It is obvious that – due to the properties of the $F-$function – values of G other than G^h, G^m, or G^l cannot be chosen in an MPE. In order to rule out $G_t = G^h$, assume that – contrary to what is stated in the lemma – there exists an MPE where $G_t = G^h$ and $\rho_t = 0$ for all t. Again, this assumption can be led to a contradiction by looking at the citizens' decision on whether or not to revolt against the elite in a certain period t: As was the case above, the assumptions about the MPE imply that condition (10) must hold, with $G_t = G^h$ on the LHS. Yet, if the condition holds with $G_t = G^h$, it must also hold with $G_t = G^m$. Thus, even if the elite deviated and chose $G_t = G^m$ in the first step (instead of setting G_t equal to the "equilibrium" value G^h), the best response would still be to abstain from ousting the ruling elite. As a result, the ruling elite must find it optimal to choose $G_t = G^m$ (i.e., to implement the level which maximizes the budget surplus). But this contradicts the initial assumption. Finally, note that assuming the existence of an MPE where – as long as $S_t = R$ – we have $G_t = G^h$ and $\rho_t = 1$ can be led to a contradiction in a similar way.

Proof of Proposition 2. Suppose that the sanctions' intensity under dictatorship is given by Λ_2 – which is exactly sufficient to induce the type-3 equilibrium. Then, the value function of the representative member of group $i \in \{1, 2\}$ is given by

$$V_i(R, \Lambda_2) = \ln\left((1 - \delta)(\Lambda_2)^{-1}(F^l - \chi_i)\right) + \beta\frac{\ln(F^h - G^h)}{1 - \beta}.$$

Using the definition of Λ_2 (equation 13) to substitute for $F^h - G^h$ and rearranging terms yields

$$V_i(R, \Lambda_2) = \ln\left(\left[(1-\delta)F^l\right]^{1/(1-\beta)} (\Lambda_2)^{-1/(1-\beta)} \frac{F^l - \chi_i}{F^l - \chi_1} \right).$$

In the absence of any sanctions, the corresponding value functions are given by

$$V_i(R, 1) = \frac{\ln\left((1-\delta)F^j\right)}{1-\beta} = \ln\left(\left[(1-\delta)F^j\right]^{1/(1-\beta)} \right),$$

whereas $j = m$ if $\Lambda_1 > 1$ (and $j = l$ otherwise). Subtracting the latter value function from the former results then in

$$V_i(R, \Lambda_2) - V_i(R, 1) = \ln\left(\left[\frac{F^l/F^j}{\Lambda_2}\right]^{1/(1-\beta)} \frac{F^l - \chi_i}{F^l - \chi_1} \right),$$

from which (15) immediately follows.

Figures

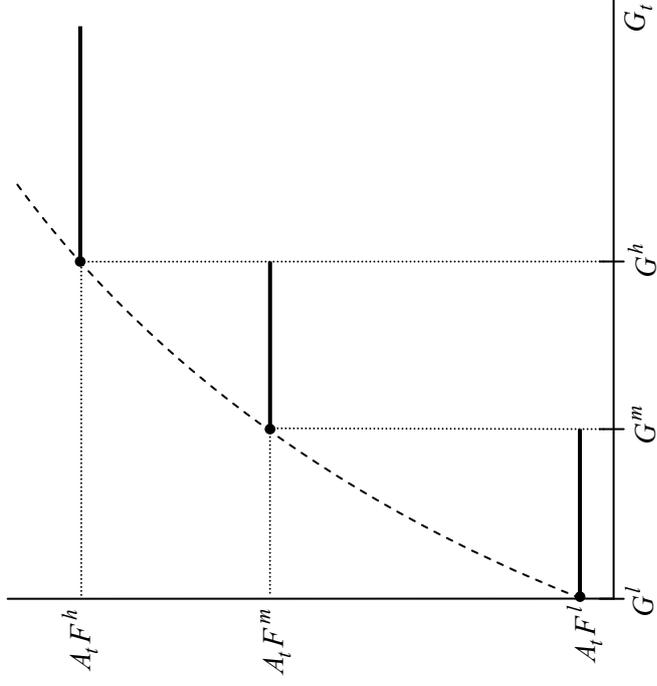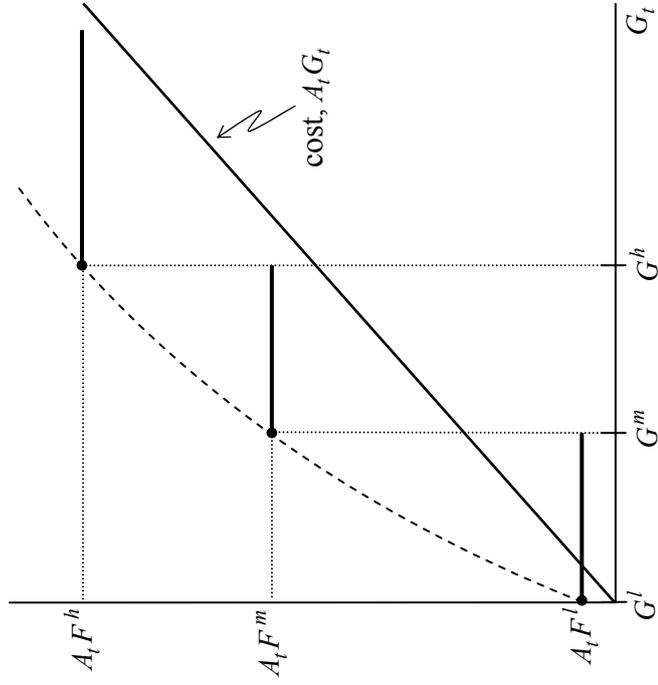*Figure* **1** – Public investment and the representative citizen's income (with $\rho_t = 0$)

Figure 2 – Public investment, private-sector income, and tax revenues (with $\rho_t = 0$)

a. Private-sector income and public expenses

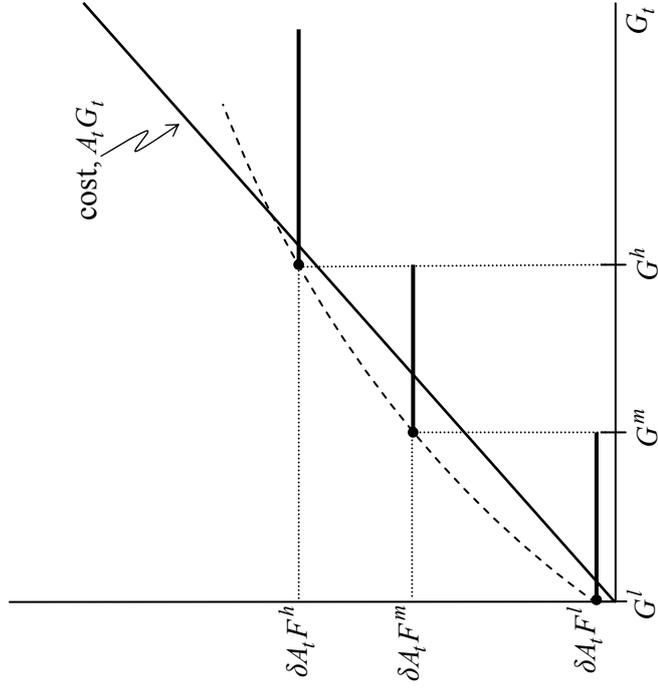*b.* Maximum tax revenues and public expenses

Figure 3 – Sanctions' intensity and the nature of the equilibrium

,,high''

,,intermediate''

,,low''

$(F^l / F^m)\Lambda_1$

Equ. 3

Equu. 2

Equ. 1

1 Λ_1 Λ_2

sanctions' intensity, Δ

Equ. 1: $\Pi_t = (\delta, G^m)$ and $\rho_t = 0$ for all t Equu. 2: $\Pi_t = (\delta, G^l)$ and $\rho_t = 0$ for all t

Equ. 3: $\Pi_t = (\delta, G^l)$ and $\rho_t = 1$ as long as $S_t = R$

***Figure* 4** – Sanctions' intensity and effectiveness (given $S_t = R$)

a. Sanctions' intensity and the political state in $t+1$
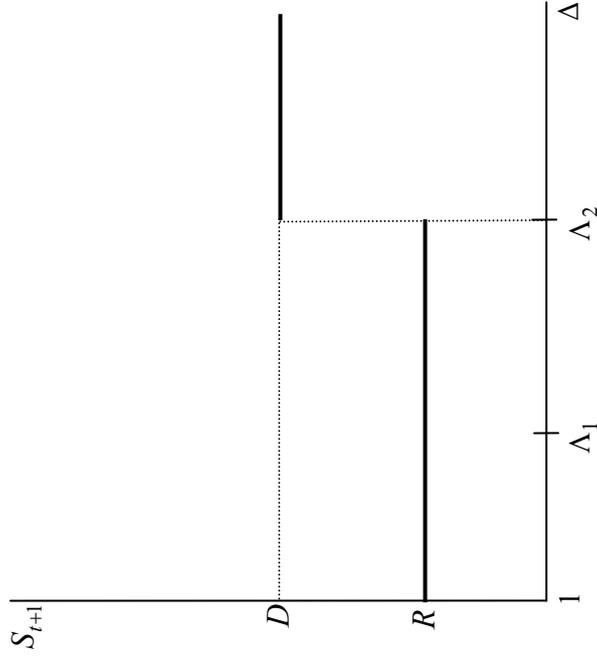
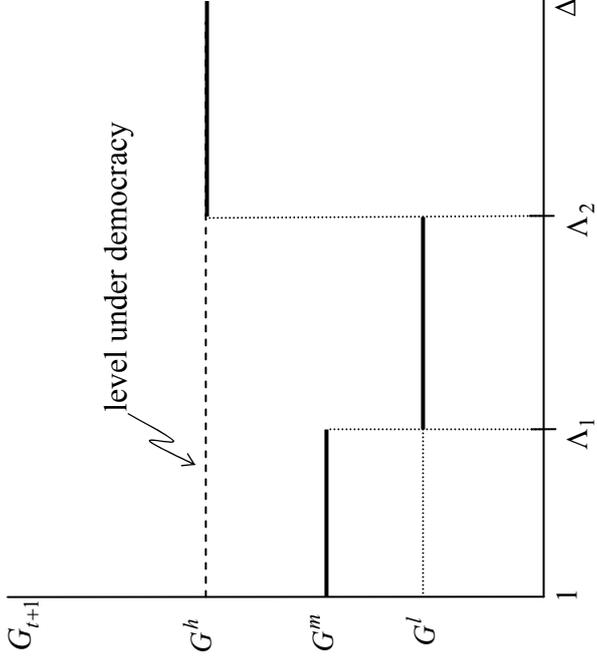*b.* Sanctions' intensity and public investment in $t+1$

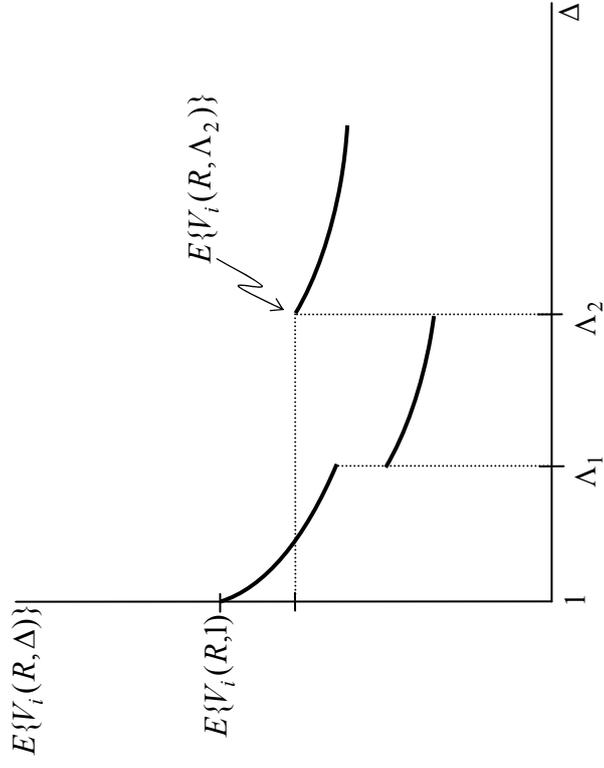*Figure 5* – Sanctions' intensity and the average payoff

a. Type-1 equilibrium in absence of sanctions

b. Type-2 equilibrium in absence of sanctions