On the Origin of States:
Stationary Bandits and Taxation in Eastern Congo

Raul Sanchez de la Sierra*
Columbia University

JOB MARKET PAPER

January 25, 2014

Abstract

The state is among the greatest developments in human history and a precursor of economic growth. Why do states arise, and when do they fail to arise? A dominant view across disciplines is that states arise when violent actors impose a “monopoly of violence” in order to extract taxes. One key fact underlies all existing studies: no census exists prior to the state. In this paper, I provide the first econometric evidence on the determinants of state formation. As a foundation for this study, I conducted fieldwork in stateless areas of Eastern Congo, managing a team that collected village-level panel data on current armed groups. I develop a model that introduces optimal taxation theory to the decision of armed groups to form states, and argue that the returns to such decision hinge on their ability to tax the local population. A sharp, exogenous rise in the price of a bulky commodity used in the video-game industry, coltan, leads armed groups to impose a “monopoly of violence” in coltan villages. A later increase in the price of gold, easier to conceal and hence more difficult to tax, does not. Results based on two alternative identification strategies are also consistent with the model. The findings support the hypothesis that the expected revenue from taxation, in particular tax base elasticity, is a determinant of state formation.

* Columbia University, Department of Economics, rs2861@columbia.edu. I am very grateful for invaluable guidance and support from Christopher Blattman, Pierre-Andre Chiappori, Macartan Humphreys, Suresh Naidu, Bernard Salanie, and Eric Verhoogen. I am grateful to Ritam Chaurey, Donald Davis, Jonas Hjort, for invaluable contributions. I am indebted to Gauthier Marchais, with whom I led the data collection. This project was supported by the National Science Foundation, the Private Enterprise Development in Low-Income Countries exploratory grants (PEDL), the International Peace Research Association Foundation (IPRAF), the Center for the Study of Development Strategies at Columbia University, the Program for Economic Research at Columbia University, the Earth Institute, the Advanced Consortium in Cooperation and Complexity and the Leitner Family Student fellowship. I thank Massimo Morelli and Peter van der Windt for their support, which was crucial in the first stages of this project. This paper draws on 18 months of fieldwork by a team of 11 surveyors, led by Jean-Paul Zibika, Christian Mastaki, and Eustache Kulumbwa.
1 Introduction

Why do states arise? The emergence of the state is among the most important developments in human history, and precedes all recorded episodes of sustained economic growth. Social scientists across disciplines have long recognized the impact of state formation on economic and human development and commonly refer to it as the “Great Transformation” (Bates, 2011). Without states to protect property rights, there are few incentives to invest, and societies will struggle to achieve growth (Grossman, 1999, Hirshleifer, 1995). Without states, there can also be disastrous human outcomes. It is precisely in a “failed state”, the Democratic Republic of Congo, where more than 5 million persons died as a consequence of the arbitrary use of violence by armed groups.2

War makes states.3 Historical accounts of state formation trace its origins to territorial conquests by medieval lords. The dominant view is that states arise when armed organizations, lords and bandits, crush local resistance and emerge with a “monopoly of violence” in their locality to collect taxes and finance wars of defense (Tilly, 1985). Supporting the criminal origins of states, Olson (1993) developed a theory of states as “stationary bandits”. Stationary bandits own the monopoly of violence, and thus promote growth in order to tax it, in contrast to “roving bandits”, who engage in uncoordinated expropriations and discourage economic activity. A fundamental fact in all disciplines is that prior to the state, there is no census. There is therefore no formal statistical evidence on the process, let alone the causes of state origins.

In this paper, I provide the first econometric evidence on the determinants of state formation. I develop a theory that integrates state formation by armed groups with the theory of optimal taxation. If medieval lords formed states in order to collect taxes, the choice to form a state must depend on their expected ability to tax. The key insight of the model is that an increase in the value of output increases the returns to form a state, only when output can be taxed. I test this theory using a novel dataset on armed groups in 140 villages of Eastern Democratic Republic of Congo. My results suggest that a positive demand shock on a bulky mineral, coltan, leads armed groups to acquire villages endowed with coltan. A later demand shock on gold, easier to conceal and hence more difficult to tax, does not have this effect in villages endowed with gold.

As a foundation for this study, I managed a team of data collectors in stateless and war-torn areas of Eastern Democratic Republic of Congo in order to collect a village-level panel dataset of yearly activity of armed groups. The data allows me to identify types of violence (territorial conquests and pillage of assets), transitions of armed groups in and out of monopolies of violence at the village level, as well as all types of taxes levied by armed groups on economic activities in the village since 1995.

---

1 “Throughout history the responses of human societies to the problems of distributing property and of allocating resources between productive and appropriative activities probably have had greater consequences for welfare than have their responses to the problem of allocating resources among different productive activities taking property as given, which is the problem on which economic analysis traditionally has focused.” (Grossman, 1994)


3 “The trimmed-down argument stresses the interdependence of war making and state making and the analogy between both of those processes and what, when less successful and smaller in scale, we call organized crime. War makes states, I shall claim.” (Tilly, 1985)
Eastern Congo is a well-suited environment to examine the formation of states. The state is considered one of the world’s two weakest states, as well as a “failed state”; its authority is challenged by armed groups who have proliferated in the East. Armed groups fight to control territory, in which they administer “monopolies of violence”, collect taxes, provide public goods, and even enjoy popular support, exactly as in historical accounts of European state formation.

In the first main result of the paper, I find that armed groups respond to an increase in the price of coltan by conquering coltan villages, where they establish a monopoly of violence and collect taxes. To establish a causal relationship I exploit a drastic demand shock for coltan. In the year 2000, innovations in the video-games industry led the demand for columbite-tantalum (coltan) to skyrocket. The US price of coltan rose abruptly from 90 US$ per kilogram to 590 US$ per kilogram at the start of the year, and collapsed at the end of the year.

In the second main result of the paper, I find that armed groups do not fight over the control of gold villages in response to an increase in the price of gold. Following September 11 and the resulting 2001-2002 global recession, a rush for gold as a safe haven induced the price of gold to rise sharply.

Having established the main results, I implement three additional tests that support my theory. i) I turn to the patterns of taxation, and find that armed groups taxes are consistent with testable implications from optimal taxation theory. They tax coltan output, but not gold output, and offer elaborate tax contracts that reflect a concern with maximizing tax revenues. ii) I then exploit the timing and targeting of a large military intervention to establish that when their time horizon decreases, settled armed groups revert to pillaging their own population. In 2009, the Congolese Army led a large-scale operation aimed at eradicating the Forces de Libération du Rwanda (FDLR), which drastically reduced their time-horizon. In response to intervention, the FDLR stopped taxing and increased violent pillaging in the villages they controlled, compared to other villages. iii) Finally, I exploit the timing and targeting of a national peace agreement to establish that settled armed groups may have a positive impact on growth. In 2003 the Rassemblement Congolais pour la Démocratie (RCD) agreed to vacate the villages under its control, leaving these villages unprotected. I find that the withdrawal of the RCD led to a decrease in economic activity in the villages it formerly controled, compared to similar villages. This suggests that settled armed groups may have a positive effect on output, consistent with the theory.

This paper’s findings have important implications for theory and policy. The findings provide an economic explanation for the origin of states and their relationship with the economy. They expand a large body of work in economics that emphasizes the role of states and institutions on economic growth and explains their evolution based on economic calculations (Acemoglu and Robinson, 2006, Acemoglu, Ticchi, and Vindigni, 2006, Bates, Greif, and Singh, 2002, Besley and Persson, 2009, Grossman, 1997, North and Weingast, 1989). This literature is largely historical and theoretical. This paper provides the first econometric evidence on the process and causes of state formation.

---


My findings also contribute to existing scholarship outside of economics. The determinants of state formation have been studied across disciplines, but are unexplored with formal statistical methods, since the formation of states takes precedence data collection capacity. Voluntaristic theories argue that states emerged as a result of a mutually beneficial social contract (Hobbes, 1651, Rousseau, 1762). On the other hand, conflict theories of state formation, which are the dominant view, compare the origin of states to organized crime. They suggest that states imposed monopolies of violence by coercion during episodes of war and conquest (Carneiro, 1970, Olson, 1993, Tilly, 1985). My data allows me to go a step further. I use optimal taxation theory to explain why organizations of violence would form states, and then test it using formal statistical methods.

I also find that price shocks give rise to local states particularly around trade routes. Historical accounts of European state formation suggest that the growth of trade allowed states to develop (Ardent, 1975). Since much of trade is observable, it can readily be taxed. My results enrich this strand of literature with disaggregated econometric evidence. In a related argument, population mobility weakens the ability of states to collect taxes and hence to arise (Carneiro, 1970, Herbst, 2011, Scott, 2009). This paper generalizes the problem of factors’ mobility into a taxation problem, where the tax authorities, armed groups, face tax bases with heterogeneous ability to avoid taxes. The findings suggest that the ability of gold producers to hide output is a key barrier to taxation by armed groups. Therefore, in this paper, I test and refine existing theories of state formation in anthropology, sociology, political science, and economics.

While the state-formation literature takes organized violence as given, the civil war literature explains the motives that lead individuals to engage in violence (Collier and Hoeffler, 2004, Humphreys and Weinstein, 2006, Weinstein, 2007). In this paper, I link these two strands of literature by providing an explanation for why actors that engage in violence may organize violence in a particular location, leading to proto-states, and vice-versa. By the same token, this paper contributes to the civil-war literature. I collected a large, detailed dataset on armed groups’ behavior. Blattman and Miguel (2010) conclude that the study of civil war is limited by the absence of high quality disaggregated data. This project, thus, expands the empirical basis of the civil war literature. The paper also contributes to the study of civil war with its findings. I find that demand for labor intensive commodities increases violence, aimed at establishing taxation monopolies. However, Dube and Vargas (2013) find that a rise in the price of a labor intensive commodity increases the returns to productive activities, hence decreases participation in violence. This paper, therefore, introduces taxation to the set of armed groups’ choices previously considered, which has implications for explaining violence as well as the types of violence. This paper’s findings, by showing that the ability of armed groups to tax reduces violence against civilians, also complement Weinstein (2007), who proposes a theory where the ability to recruit soldiers determines violence. Last, this paper also contributes to the study of civil war by its identification strategy. To establish causality, I exploit large price shocks and political events. This strategy compares to an emerging civil war literature (Besley and Persson, 2008, Dube and Naidu, 2010, Dube and Vargas, 2013, Nunn and Qian, 2012).

---

6The literature on state formation is separated between early and modern state formation. The literature on early state formation is reviewed in Claessen and Van de Velde (1991) and Claessen and Skalnik (2011).
This paper also contributes to the understanding of the relationship between resources and government accountability (Bates and Lien, 1985). One strand of literature addressing this link is the resource curse literature.\textsuperscript{7} When states have access to valuable resources, they will be less dependent on the population for taxes and hence will be under less pressure to develop participatory institutions or provide public goods. In response to positive resource shocks, I find that armed groups develop taxation monopolies focused on taxing mining activity, but also other economic activity. My results hence expand this literature by suggesting that if the value of resources can only be extracted through taxes on labor, they may not have a detrimental effects on institutions.

The remainder of the paper is organized as follows. Section 2 presents the setting and Section 3 presents the theory. Section 4 describes the data and data collection strategy. Section 5 examines the effect of price shocks on armed groups’ acquisition and formation of monopolies of violence. In Section 6, I examine the relationship between armed groups and the economy. Section 7 provides a brief discussion and Section 8 concludes.

2 Background

Owing to the collapse of the central state, armed groups in the Democratic Republic of Congo have proliferated since the 1990’s.\textsuperscript{8} To finance their operations, they collect taxes on the mineral sector and other activities. Currently, there are 39 identified armed groups.

2.1 The two phases of the conflict


The Second Congo War (1998-2003) involved a large number of armed groups, and is referred to as the “Great African War”. In 1998, the Rassemblement Congolais pour la Démocratie (RCD) launched an offensive to overthrow the former DRC president in office, Laurent-Désiré Kabila.\textsuperscript{9} The coup did not succeed, but a myriad of RCD units sustained the control of the Eastern provinces, while the Congolese state defended the West. The RCD struggled to dominate the rural areas, where it faced resistance by the self-defense groups known as the Mayi-Mayi and by the Forces De Libération du Rwanda (FDLR). These groups formed temporary alliances, but they were internally divided. The war also involved the participation of nine foreign armies and thirty local militia, who fought mostly in the east of the country.

Despite the support from the UN Peacekeeping force (MONUSCO), the “post-conflict” period (2003-2013) is characterized by the struggle of the Congolese state to regain control over the Eastern regions. Following a peace agreement signed in Sun-City (South-Africa) in 2003, the Rassemblement Congolais pour la Démocratie (RCD) agreed to vacate the East of the Congo and integrate the

\textsuperscript{7}See Bannon and Collier (2003) and Humphreys, Sachs, and Stiglitz (2007) for reviews of the literature.
\textsuperscript{8}The Democratic Republic of Congo was named Zaire until 1997.
\textsuperscript{9}The First Congo War lasted from 1996 to 1997. It ended with the overthrow of President Mobutu.
national government. The subsequent withdrawal of the RCD in 2003 and 2004 left large areas in the East without protection. In the immediate aftermath, the Forces De Libération du Rwanda (FDLR) and local self defense groups increased their territorial control. Only in 2009, the Congolese army, together with the United Nations peacekeeping force and Rwanda, led a major operation against the FDLR. This operation, named Kimia II, successfully weakened the link between the FLDR and its former tax base, although it failed to eradicate it. Armed groups continue to operate in the East, where they control up to 95% of the territory in some administrative divisions. In May 2012 a newly formed armed group, the M23, threatened to gain control of the Eastern regional capitals, and established its own administration in a large territory, which included numerous Ministries. The Congolese army and the United Nations ultimately eradicated it in November 2013. Ideology, self-defense, as well as economic interests animate these groups. Economic motives, however, stand out as one of the major motivations across the two periods.

2.2 The mining sector

Armed groups developed economic interests during the Second Congo War, which persisted during the post-conflict period. One of their major source of revenues is the taxation of mineral trade. Extraction of minerals in Eastern Congo is labor intensive, and is a major livelihood of the local population. World Bank (2008) estimates that there are between 500,000 and 2 millions artisanal miners in DRC, responsible for 90% of DRC mineral production. Heavy minerals, with high volume productivity of labor (coltan, cassiterite), coexist with light minerals (gold), which are harder to find and, thus, to produce. So-called artisanal mining requires minimal capital and skills. After miners have extracted the heavy minerals, they hand them over to carriers, usually in bags of 50 kilograms. Carriers walk up to one day to the closest village (support village), where local traders purchase the output. Owing to the large volumes produced and a poor road infrastructure, traders ship heavy minerals to Bukavu and Goma by plane. Since trade is vulnerable to pillage, armed groups provide security and collect taxes along the carrier’s route. They play a central role in the protection of business (Verweijen, 2013). Gold output, however, is easy to conceal. Because of its large value to weight ratio, gold miners conceal the gold output they are able to find. Miners and traders smuggle gold directly through Burundi, Bukavu or Uvira, circumventing taxes, and Congolese border customs. According to World Bank (2008) estimates, the value of gold exports is $125 million, most of which is smuggled. A central feature across minerals is that armed groups provide security and collect taxes at the mine as well as on other economic activity at the village.

\[10\] In Shabunda, an administrative division in Sud-Kivu that this survey visited, the Raia-Mutombokis control 95% of the territory. See for instance http://radiookapi.net/actualite/2013/02/28/shabunda-la-milice-raia-mutomboki-occupe-95-du-territoire-selon-son-administrateur/


\[12\] The average world price per kilogram of gold in the period was US$17,404, against US$136 for coltan. Daily production per worker is approximately 20 kg in coltan, and between 1 and 10 grams in gold. See De Failly (2001), De Koeing (2009), Nest (2011), Geenen (2013) and Vlassenroot and Raeymaekers (2004) for descriptions of the mineral sector.
De jure, property rights are poorly defined. Land belongs to village chiefs according to customary law, who then lease it out to individuals.\textsuperscript{13} Formal law, however, often contradicts traditional law. A large number of state or quasi-state agencies, including criminal groups, tax miners’ activity in seemingly arbitrary ways. De facto, however, the production process is highly organized, with a well-defined hierarchy. Mining workers are usually residual claimants. Taxes are anticipated, and armed groups often cooperate with the population by offering protection in exchange for taxes. Since production is labor intensive, armed groups rarely mine themselves, but tax miners instead.

In the period for which I have data, the mining sector is marked by two price shocks. First, at the start of 2000, Playstation, a subsidiary of Sony, announced the release for Christmas of a new promising product, which processed columbite-tantalum (coltan). Since coltan extraction, mostly done in Australia by the year 2000, required large investments, supply was inelastic. In response to the announcement, coltan processing firms rushed for coltan from other areas where supply could respond, leading the daily price of coltan to skyrocket from $90 per kilogram to $590 per kilogram, where it persisted until year end. The DRC emerged as a major substitute of Australia’s coltan. Second, following 9/11 and the economic downturn, investors rushed for safe assets, including for gold. This led the price on gold to rise in the aftermath of the crisis of 2002, which continued to rise during the post-conflict period (2004-2013). Figure 7 shows the world prices of gold and coltan.\textsuperscript{14}

3 A theory of stationary bandits

In this section, I develop a model where an armed group (henceforth, the bandit) can acquire resources by taxing households’ activities. The objective of the model is to elicit how the bandit’s ability to tax determines the returns to acquiring monopolies of violence in order to tax. Empirically, the main causal effect of interest is the effect of mineral demand on armed groups choice to establish or quit taxation monopolies. The model generates four insights. First, an increase in global demand for minerals that are easy to tax (coltan) increases the returns to acquire a monopoly of violence ($P_1$). Second, an increase in global demand increases the returns of taxation monopolies especially in areas where output can be traded ($P_1'$). Third, an increase in global demand for minerals that can easily be concealed, and hence harder to tax, results in a weaker effect on the returns to acquire taxation monopolies ($P_2$). I derive three other sets of testable implications the model.

i) Since armed groups acquire resources through taxation, taxes should be inversely proportional to their respective tax base elasticities; in particular, since gold is easy to hide, the tax on gold output should be very small, or absent; ii) The revenues to engage in uncoordinated expropriation of civilian assets increase when the bandit’s time horizon shrinks; iii) When they own a monopoly of violence, bandits may have a positive effect on economic activity. The model applies established results of optimal taxation and contract theory to the decision to form proto-states. The model is similar to Besley and Persson (2013), although it ignores dynamics. The tax base elasticity can refer

\textsuperscript{13}The Congolese legal system comprises “customary” law, which is the law inherited from local customs. According to customary law, the village chief is the ultimate authority in the village and owner of the land.

to the population mobility across villages, which has played a central role in historical accounts of state development in Africa (Herbst, 2011), and East Asia (Scott, 2009). Tax base elasticity also captures general mobility of factors of production, emphasized in (Bates, Ndulu, O’Connell, Collier, and Soludo, 2008) and (Bates and Lien, 1985).

Anecdotal observation dictates the following modeling choices. First, the model considers taxes on inputs and outputs. Armed groups indeed levy taxes on mineral output, taxes on mining labor inputs, taxes on food sales, taxes on transit of persons, taxes on village mills and poll taxes. I focus on labor taxes and output taxes for parsimony, and because they capture in essence the behavioral responses in all other activities. Second, the model considers a static framework where bandits can choose between acquiring civilian assets through uncoordinated pillage or instead attempt to impose a “monopoly of violence”, in which case they may extract revenue through anticipated taxation. These extremes also reflect the anecdotal observation of armed groups’ behavior in Eastern Congo. Of 600 recorded violent events in the sample villages, 56% are pillage operations, aimed at capturing village assets. On the contrary, 40% are conquest attempts, aimed at gaining the monopoly of violence in a village, by defeating a competing armed group. Once established, armed groups collect taxes. Figure 1 shows that conquests take place early in the morning, possibly to surprise the defense force, and pillage operations take place at sunset, consistent with a crime deterrence hypothesis (Becker, 1968). This supports the empirical distinction between the two types of expropriation. Finally, I ignore strategic interactions, because in the period of the coltan shock, one armed group dominates the rest militarily. I discuss limitations of doing so in Section 7.3.

3.1 Model setup

Consider one village economy and a mass of roving bandits. Roving bandits engage in uncoordinated expropriation of households’ assets and output. They are roving because they have a short time horizon, and do not internalize future production in the village. One bandit, however, has superior force, and can choose to impose his “monopoly of violence” in the village, turning into a stationary bandit. If he acts as stationary bandit, he announces his theft plan to households (a tax plan) so as to maximize his taxation revenue. The village is composed of k of identical households, j = 1, 2, ..., k. In what follows, I drop the household identifiers. Households are endowed with assets and choose the level of the following variables in N sectors i = 1, 2, ..., n: labor supply, the amount of labor to hide from the tax on labor, and the volumes of output to hide from the tax on output. The stationary bandit is a Stackelberg leader and households are uncoordinated followers.

At step 0, the strong bandit chooses whether to compete over village assets by means of pillaging (as a roving bandit) or to acquire the monopoly of expropriation at the village (become a stationary bandit), at a fixed cost F. The fixed cost captures the costs incurred in the attempt to conquer the village (waging troops, logistics, and expected losses for the group), the investments made to maintain the control of the village, and potentially, the costs of administering the village. At step 1, the stationary bandit expropriates. If, in step 0, the strong bandit chooses to acquire resources by pillaging, he competes with other roving bandits for expropriation. For simplicity,
I assume that roving bandits, since they do not have a long term relationship with households, can’t commit not to fully expropriate village assets and output. When there are only roving bandits, therefore, households anticipate that they will be expropriated with probability 1 and do not produce. Therefore, roving bandits can only expect to acquire household assets. If, in step 0, the strong bandit chooses instead to acquire the monopoly of violence (and becomes a stationary bandit), I assume he is able to develop a relationship with the households that allows him to commit to his promise of theft.\textsuperscript{15} The stationary bandit expropriates using taxes on output $\tau = (\tau_1, \tau_2, \ldots, \tau_n)$, taxes on labor inputs $t = (t_1, t_2, \ldots, t_n)$ and potentially poll taxes.\textsuperscript{16} He announces how much he plans to steal (tax) such as to maximize his revenues taking into account behavioral responses by the households. I first present the households’ and stationary bandit’s problems in step 1, when the strong bandit has acquired the monopoly of violence. I then derive comparative statics of the effect of price shocks on the bandit’s choice to acquire a monopoly of violence. In a second part, I derive additional testable implications reflecting the relationship the stationary bandit and the economy. I focus on the three sets of implications: i) properties of optimal taxes; ii) effect of the time horizon; iii) economic impact of the stationary bandit.

### 3.2 The stationary bandit’s economy

This section lays out the optimization problems of households and the stationary bandit, and discusses under which conditions it will not be possible to tax gold output.

#### 3.2.1 Optimization problems of the households and the stationary bandit

Labor volume productivity, $\tilde{\alpha} \in R^n$, is stochastic, and has the cumulative distribution function $F(\tilde{\alpha})$. The sector volume productivities, $\tilde{\alpha}_i \in R$ are independent across sectors, and have a known cumulative distribution function $F_i(\tilde{\alpha}_i)$. Output in sector $i$, $\tilde{y}_i$ is a linear function of labor inputs and volume productivity: $\tilde{y}_i = \tilde{\alpha}_i e_i$. Households are residual claimants. They sell the output they produce in the world markets, where the prices of outputs, $p_i$, are exogenous. Households’ Bernoulli utility is concave in the unique consumption good, and the disutility from labor, $c(e)$, is separable across sectors as well as convex in each sector labor supply.

Prior to the realization of uncertainty, but knowing the distribution of volume labor productivities and the vectors of taxes and prices, the households choose how much labor to supply to each sector $e_i$, $i = 1, 2, \ldots, n$ and how much labor to hide from labor taxes $e_i^H$, $i = 1, 2, \ldots, n$. Then, upon observing realized production, households choose the volume of output they conceal from taxes.

\textsuperscript{15}The assumption of binding commitment by the stationary bandit reflects features of a repeated relationship between the bandit and households. Without this assumption, the bandit will never be able to commit and will always expropriate 100% after production. Anticipating this, the households would not produce. However, this scenario leads back to anarchy. Analogous behavior can be obtained by simply introducing a cost in the bandit’s payoff from deviating from the taxation plan. For similar treatment of the future, see Dixit (2004).

\textsuperscript{16}Poll taxes are usually defined as taxes on persons. They are especially easy to administer, since they require no prior observation of economic activity or estimation of income, and are conditioned on fixed households. I do not provide a formal treatment here.
Households’ revenues from production in sector $i$ take the following form:

$$\tilde{\pi}_i = (1 - \tau_i)p_i\tilde{\alpha}_i e_i + \tau_i p_i H_i - t_i (e_i - e_i^H) - G^i(H_i, p_i) - E^i(e_i^H, p_i)$$

where $G^i(H_i, p_i)$ and $E^i(e_i^H, p_i)$ are respectively the costs of concealing output and hiding labor. Therefore, $\tau_i p_i H_i - G^i(H_i, p_i)$ and $t_i e_i^H - E^i(e_i^H, p_i)$ are respectively the profits from concealing output and hiding labor. I assume the two cost functions to be differentiable, monotonically increasing in both arguments, $(G_1^i > 0, E_1^i > 0, G_2^i > 0, E_2^i > 0)$ and strictly convex $(E_{11}^i > 0, G_{11}^i > 0)$. In what follows, I focus on solutions where $e_i \geq 0$, $H_i \geq 0$. The household’s problem is:

$$\max_{e : H \in H} \int u \left( \sum_{i=1}^N \left( (1 - \tau_i)p_i\tilde{\alpha}_i e_i + \tau_i p_i H_i - t_i (e_i - e_i^H) - G^i(H_i, p_i) - E^i(e_i^H, p_i) \right) \right) \ dF(\tilde{\alpha}) - c(e)$$

s.t. $H_i \leq \tilde{\alpha}_i e_i$, $i = 1, 2, \ldots , N$

$e_i^H \leq e_i$, $i = 1, 2, \ldots , N$

$\sum_{i=1}^N e_i^3 = L$

where $L$ is the household’s total time endowment, and $\tilde{\alpha} \in \mathbb{R}^n$ is the vector of labor productivity realizations. The problem is akin to a portfolio allocation problem, albeit here the household allocates its labor to sectors with uncertain volume productivities $\tilde{\alpha}_i$; investment is costly (the disutility of labor supply); investment is taxed (the taxes on labor inputs); the returns to investment are taxed (the taxes on output); and the household can conceal part of the returns to investment (by concealing output). The household’s first order condition implies conditions under which the household will conceal the totality of output produced. For a given combination of taxes, prices, labor supply, and volume productivity realization, the household will conceal all output produced if and only if the marginal revenue from concealing an additional unit exceeds its marginal cost, i.e. when $\tau_i p_i > G_1^i(\tilde{\alpha}_i e_i, p_i)$. In that case, the optimal volume concealed is a function of $\tau_i$, $p_i$, $e_i$, but also of the productivity realization $\tilde{\alpha}_i$, i.e. $H^*_i = H_i(p_i, \tilde{\alpha}_i e_i, \tau_i)$. However, when the household conceals only part of the output he produced (at interior solutions), the volume concealed is independent on the realization, i.e. $H^*_i = H_i(p_i, \tau_i)$. The volume concealed, as a function of realized output has the shape represented in Figure 2. The household’s optimal choice is characterized by the following functions: $e^*_i(p, F_\alpha, \tau, t)$, $e_i^{H*}(p_i, t_i)$, $H^*_i(p_i, \tilde{\alpha}_i e_i, \tau_i)$. The sector $i$ labor supply, $e^*_i$, is decreasing in $p_j$, $\forall j \neq i$. In the absence of income effects, it is increasing in $p_i$ and in $E\alpha_i$. Finally, with sufficient prudence (large $u'''$), $e^*_i$ is decreasing in $\Var(\alpha_j)$.

The stationary bandit chooses taxes to maximize his tax revenue. Since taxes are communicated to households before they make their choices, the stationary bandit internalizes the households’ behavioral responses at the time of choosing the taxes. The stationary bandit’s problem is:

\[\text{max } \int u \left( \sum_{i=1}^N \right) \ dF(\tilde{\alpha}) - c(e)\]

s.t. $H_i \leq \tilde{\alpha}_i e_i$, $i = 1, 2, \ldots , N$

$e_i^H \leq e_i$, $i = 1, 2, \ldots , N$

$\sum_{i=1}^N e_i^3 = L$

\[e^*_i(p, F_\alpha, \tau, t)\]

\[e_i^{H*}(p_i, t_i)\]

\[H^*_i(p_i, \tilde{\alpha}_i e_i, \tau_i)\]

\[\text{max } \int u''' \left( \sum_{i=1}^N \right) \ dF(\tilde{\alpha}) - c(e)\]

s.t. $H_i \leq \tilde{\alpha}_i e_i$, $i = 1, 2, \ldots , N$

$e_i^H \leq e_i$, $i = 1, 2, \ldots , N$

$\sum_{i=1}^N e_i^3 = L$

\[e^*_i(p, F_\alpha, \tau, t)\]

\[e_i^{H*}(p_i, t_i)\]

\[H^*_i(p_i, \tilde{\alpha}_i e_i, \tau_i)\]

\[\text{max } \int u''' \left( \sum_{i=1}^N \right) \ dF(\tilde{\alpha}) - c(e)\]

s.t. $H_i \leq \tilde{\alpha}_i e_i$, $i = 1, 2, \ldots , N$

$e_i^H \leq e_i$, $i = 1, 2, \ldots , N$

$\sum_{i=1}^N e_i^3 = L$
\[
\max_{\tau, t} \sum_{i=1}^{N} \left[ p_i \tau_i \left( \alpha_i e_i^* (p, F, \tau, t) - H_i^* (p_i, \tilde{\alpha}_i e_i, \tau_i) \right) + t \left( e_i^* (p, F, \tau, t) - e_i^{\text{H}} (p_i, t_i) \right) \right]
\]

When input taxes are not available, the revenue maximizing tax rates on output (optimal tax) satisfy the following relationship:

\[
\frac{\tau_i^*}{1 - \tau_i^*} = \frac{1}{\tilde{\varepsilon}_i^Y} - \sum_{i \neq l}^{N} \frac{\tau_i p_i Y_i \varepsilon_i^Y}{\tilde{\varepsilon}_i^Y} \tag{1}
\]

where:

\[
\tilde{\varepsilon}_i^Y = \frac{\partial \tilde{Y}_i}{\partial (1 - \tau_i)} (1 - \tau_i)
\]

is the elasticity of observable output in the sector and \( \tilde{Y}_i = (\alpha_i e_i - H_i) p_i \). Let

\[
s_i^e = \frac{\alpha_i e_i}{\alpha_i e_i - H_i}, \quad s_i^H = \frac{H_i}{\alpha_i e_i - H_i}
\]

The elasticity of observable output \( \tilde{\varepsilon}_i^Y \), can be rewritten as \( s_i^e \tilde{\varepsilon}_i^e + s_i^H \tilde{\varepsilon}_i^H \). This formula captures that the optimal tax is proportional to the inverse of factor elasticities, a well-established result of optimal taxation (Ramsey, 1927). In the absence of distributional concerns, the indexes of relative discouragement do not enter the optimal tax formula. The higher the elasticity of the sector’s taxable output, the smaller the optimal tax rate. The last term in the denominator captures the fiscal externalities arising in the presence of multiple taxable activities. Increasing the tax rate in sector \( l \) might induce reallocation of labor to other sectors raising the optimal tax.

### 3.2.2 The impossibility to tax gold output

While a worker in coltan mines might produce up to 50 kilograms of coltan per day, a typical day of work in gold yields between 0.1 and 10 grams of gold output. In the absence of sophisticated monitoring technology, concealing gold output must be costless (Geenen, 2013). This feature of gold production is reflected on this project’s qualitative fieldwork. Furthermore, anecdotal evidence on gold mines elsewhere, shows that employers frequently use X-rays to monitor gold miners and prevent theft.\textsuperscript{19} The majority of villages in Eastern Congo is not electrified, and importing X-rays is unlikely to be affordable. Armed groups’ inability to monitor gold output is therefore an important characteristic affecting the choice whether to become the stationary bandit at gold mines. This resonates with the local say that gold is “immaterial”, because it can never be observed.

\textsuperscript{19} As part of the mine’s security, workers at the end of their shift are ushered into a corridor surrounded by glass and monitored by video cameras. Security guards carefully pick through the workers clothes and give them random full body X-ray.” See: http://factsanddetails.com/world.php?itemid=1235 & subcatid=324.
Assumption G1: Let $H_i \in R$ such that $\forall H_i < H_i$, $G_i^1(H_i, p_i) = 0$. Let $\bar{\pi}_g = \sup \{ A_g \}$, where $A_g$ is the set of gold volume labor productivity realizations. I assume that: $\bar{\pi}_g L \leq H_i$.

Assumption G2: Let $F_i$ be the fixed cost of levying an output tax on sector $i$. I assume that the expected volume productivity of labor in gold, $E\tilde{\alpha}_g$, is sufficiently low, so that the expected returns from taxing output are lower than the fixed cost of taxing output: $\tau_i^g p_g E (\tilde{\alpha}_g e_g - H_g) < F_g$.\textsuperscript{20}

Proposition 1: Under assumption G1, $\forall \tau_i > 0$, $H_i^* = \tilde{\alpha}_i e_i$. Under assumption G2, $\tau_i^* = 0$.

Assumption G1 fits the anecdotal evidence, and guarantees that it is optimal to conceal any output realization in gold. Under assumption G1, a tax on gold output is irrelevant. Assumption G2 guarantees that it is not profitable to tax gold output. Assumptions G1 and G2 are therefore both plausible and separately sufficient conditions for $\tau_i^* = 0$. This result is non-trivial: while lower average volume labor productivity $E\tilde{\alpha}_i$ leads to lower output taxes, it does not necessarily lead to absence of output taxes, as can be seen from Equation 1.\textsuperscript{21}

3.3 Main predictions: Effect of prices on emergence of violence monopolies

The strong bandit will choose to impose a monopoly of violence if:

$$\tilde{\pi} V(p, F, a_0, M, L) - F > V_r$$

Where $\tilde{\pi} = \pi/\beta$ is the effective discount factor. The term $\pi$ is the probability that the stationary bandit maintains the monopoly of violence until production realizes, and $\beta$ is the time preference. Let $i = c$ denote the coltan sector, $i = g$ denote the gold sector, and $\tilde{e} = e_i - e_i^H$ the observable labor input in sector $i$. Assume a village is endowed only with one sector (either gold or coltan), and that either assumption G1 or G2 holds. Then:

$$\frac{\partial V}{\partial p_c} = \tau_c \alpha_c \left( e_c + p_c \frac{\partial \tilde{e}_c}{\partial p_c} \right) + \frac{\partial \tilde{e}_c}{\partial p_c} t_c > 0,$$

and

$$\frac{\partial V}{\partial p_g} = \frac{\partial \tilde{e}_g}{\partial p_g} t_g > 0.$$

\textsuperscript{20}This assumes that the left hand side is increasing in $E\tilde{\alpha}_g$. From the first order conditions, it can be seen that $e_c$ is increasing in $E\tilde{\alpha}_g$ in the absence of income effects, and $H_g$ is independent on $\tilde{\alpha}_g$ at an interior solution. Finally, under the same assumptions, from equation 1, $\tau_i^*$ is increasing in $E\tilde{\alpha}_g$.

\textsuperscript{21}In particular, the optimal volume concealed $H_i^*$ is a function of $\tau_i$. Concealing all realized output for any $\tau_i$ is only possible if $\forall \tau_i > 0$: $\tau_i p_i > G_i^1(\tilde{\alpha}_i e_i, p_i)$. However, under the assumption that $G_i^1 > 0$, $\forall H_i$, there exists always some $\tau_i(\tilde{\alpha}_i e_i) > 0$, such that $\forall \tau_i < \tau_i(\tilde{\alpha}_i e_i), H_i^* < \tilde{\alpha}_i e_i$. It must therefore be that there is either a fixed cost of imposing a tax, or that the marginal cost of hiding is approximately zero at low levels of volume produced.
This expression shows that a rise in the output price increases the revenues from becoming a stationary bandit in the village through three distinct channels. First, it raises the value of output taxed, for a given level of output volume and tax level. Second, by leading to an increase in labor supply, it increases the volume of output produced, and hence the value of output taxed. Finally, since labor supply increases, the price shock also increases the revenues from taxing labor inputs. In gold villages, only the third channel changes the revenues to the bandit. The relative effects in gold and coltan depend on the labor supply responses and the level of input taxes. Estimation of the labor supply elasticities $\frac{\partial \tilde{e}_g}{\partial p_g}$, $\frac{\partial \tilde{e}_c}{\partial p_c}$, and the labor input taxes, $t_c$ and $t_g$ provides empirical support for this claim. Proposition 2 establishes sufficient conditions for this result formally.

Proposition 2: Let $G(H) = \frac{b_h}{2}H^2$, $G(e^H) = \frac{e}{2}e^H^2$, and $c(\ell) = \frac{1}{2}e$. Using empirical estimates, $\frac{\sigma^2_g}{\sigma^2_c} < 1$. Then,

$$\frac{\partial V}{\partial p_c} - \frac{\partial V}{\partial p_g} > 0$$

Furthermore,

$$\frac{\partial V}{\partial p_c} - \frac{\partial V}{\partial p_g} > 0$$

These results motivate the main empirical predictions.

**Prediction P1**: Price shocks and the returns to form states. A rise in the price of coltan increases the value of establishing a state in coltan villages. As a result, in coltan villages, armed groups will impose a village state in order to tax. In villages in which there was no state, armed groups will establish a monopoly of violence; in villages in which there was already a monopolist of violence, they will attempt to conquer it, leading to the emergence of new stationary bandits when conquests are successful. In addition, in both cases, a rise in the price of coltan should lead to investments aimed at deterring entry by competing invaders in coltan villages. In the empirical section, these investments are operationalized as acquisition of firearms.

**Prediction P1′**: Price shocks, trade infrastructure, and the returns to form states. A rise in the price of coltan increases the value of establishing a state, especially in coltan villages near trade infrastructure. Since coltan can only be shipped by plane, Prediction P1′ implies that Prediction P1 will hold especially in villages sufficiently close to airports.

**Prediction P2**: Price shocks, gold, and the returns to form states. A rise in the price of gold increases the value of establishing a state in gold villages less so than an identical value per unit increase in the price of coltan in coltan villages. As a result, armed groups will respond to a price rise in gold less than in response to an identical value per unit increase in the price of coltan.
3.4 Additional testable implications: armed groups and the economy

In this section, I first derive properties of the armed group optimal tax (i). I then discuss the effect of the time horizon (ii), and the economic effects of stationary bandits (iii).

i) It is straightforward from the inverse elasticities formula that the stationary bandit will only tax economic activity that cannot be fully hidden. Were a tax to be levied in activities that can be perfectly hidden, households would leave no taxable revenue to the stationary bandit. As a result, the stationary bandit’s taxation is limited to the following activities: observable transit, observable output, observable labor, sales of agricultural products at markets (where they can be observed), and taxes per person living in the village. Four additional predictions follow from the optimal tax.

First, when there is only one sector, and in the absence of income effects, a rise in the price of output, by decreasing the elasticity of labor supply, may increase the optimal output tax.

Second, in the presence of multiple sectors, taxes will reflect the presence of fiscal externalities:

$$cov \left( \tau_l, \sum_{i \neq l} \tau_i \bar{p}_i Y_i \right) > 0$$

This captures the role played by the presence of multiple sectors on the level of the optimal tax. If the stationary bandit collects taxes in various sectors, raising the tax in one sector leads to a reallocation of factors towards other sectors, which he may be able to tax, increasing his tax revenues from other sectors. This “fiscal externality” raises the stationary bandit’s optimal tax.

Third, the optimal tax on households (poll tax) may be positive. If the cost of migrating is high, the elasticity of households’ location is low. Since estimating household income is costly and households are observable, poll taxes were common among early Western states (Ardent, 1975).

Fourth, if possible, it is optimal for the stationary bandit to increase the households’ cost of concealing output or hiding labor. By increasing the cost of sheltering activities, bandits decrease elasticities, the optimal tax rate, and increase tax revenues. To increase the tax revenues from the transit of persons, the stationary bandit may regulate village entries, since regulation of transit decreases the elasticity of activities related to transit. To increase the tax revenues from agricultural produce, the stationary bandit may prohibit access to alternative markets to the one they tax. These actions are consistent with evidence on early modern states (Ardent, 1975).

Finally, with joint risk aversion, or when labor and outputs can be partially hidden or concealed, the stationary bandit taxing optimally one sector will offer a mix of labor input and output taxes.

ii) When the value of the future decreases, for instance from a decrease in $\bar{\pi}$, pillage becomes relatively more attractive than sustaining a monopoly of violence. A reduction in $\bar{\pi}$ can arise if the stationary bandit’s monopoly of violence in the village comes under threat.

iii) In the absence of a monopoly of violence, uncoordinated roving bandits expropriate households with certainty, leading to no production. When a bandit establishes a monopoly of violence, he chooses a taxation plan that maximizes his revenues, which will lead to production. Therefore, when a village has a stationary bandit, production is likely to be higher than otherwise.
4 The data

In this section I describe the data. I first describe how I collected the data. I then describe how I defined the main variables in the survey, and what strategies I used to minimize measurement error at the data collection phase. Following these descriptions, I present the summary statistics.

4.1 Data collection

In order to obtain data on armed groups’ violence and taxation, I designed a data collection project in Eastern Congo. I trained and monitored 11 local surveyors who collected the data between May 2012 and September 2013. In each village, the surveyors trained a group of village history specialists (village chiefs, elders, and mining sector experts) for the collection of non-sensitive historical data under adequate compensation (henceforth, the village specialists). The data collection activities followed a strict protocol, designed to last 7 days per village. During the day, the surveyors implemented retrospective household surveys, in private. In the evenings they provided guidance and corrected the progress of the village specialists. In the last day of data collection activities in each village, after having implemented 8 household surveys, the surveyors held a day-long meeting with the village specialists. In this meeting, the researchers and village specialists confronted the data gathered from the different sources, and surveyors collected additional sensitive data based on the information they gathered in the household surveys. The data from this meeting compiles information from all activities and is the main source of data used in this paper.

Within each administrative division, I drew a random sample of gold villages, and the population of coltan villages. To identify the villages endowed with gold, coltan, or other minerals, I used the data from a mapping implemented in 2009 by IPIS, a Belgian research institute, and SAESSCAM, a local mining institute. To verify these lists, I interviewed in Bukavu a large number of mineral dealers working through Sud-Kivu, during 3 months of preparatory work. Based on the compiled dataset, I endowed the surveyors with a list containing all mining villages of each administrative division. In order to strengthen the representativeness of the sample to be selected, I designed a strict prospection protocol. Following this protocol, in the first week of work in each administrative division, the surveyors verified the existence and endowments of the villages in the list, and added mining villages they discovered to exist. Upon finishing prospection work, and in areas where they had access to electricity, the surveyors communicated the cross-checked lists by satellite phone. I then implemented random sampling of gold villages using a statistical package and communicated the selection using satellite phone. In areas without access to electricity, the surveyors followed a strict protocol for selecting gold villages, based on a list of random numbers I previously generated using a statistical package.\footnote{This was only applied in one administrative unit (Shabunda).} Figure 3 maps the villages in the sample based on their endowment of minerals. The villages endowed with coltan and gold are well distributed across the various administrative divisions, with no clustering of villages by mineral endowment.
4.2 Measurement

This section describes the key variables I use in this paper.

The village mineral endowment was already identified during *prospection* work. I did not record any discrepancy between the mineral endowment from prospection and survey data indicating whether the village had exploited a given mineral. Therefore, I define the mineral endowment for this paper as whether the mineral has been exploited at any point for which I have data.

I take the yearly prices of minerals from United States Geological Survey (2010). Figure 7 shows the world prices of gold and coltan.

Based on preparatory fieldwork, I define the presence of a *stationary* bandit in a given village as whether an armed person or group had a monopoly of violence in the village for at least 6 months in a given year. *Stationary* bandits are such a common phenomenon in Eastern Congo that villagers faced no difficulties identifying *stationary* bandits from *roving* bandits; the former have no intention to leave the village in the short-term, and usually provide protection to the whole village.

To measure violent events, focus on two major types of attacks, which reflect the anecdotal evidence. I measure *conquest attempts* by whether an armed group engages in violence with another armed group, with the aim of acquiring the monopoly of violence of the village. I define *pillages* as violent events on the village aimed *exclusively* at capturing assets. This definition reflects the observation that local populations are able to distinguish the purpose of the attacks unambiguously. *Pillages* also differ from *conquest attempts* on observable characteristics, which facilitates their distinction for local populations. In a *pillage*, armed groups usually arrive at sunset and flee with the village assets within a narrow period of time. *Conquest attempts* take place early in the morning and violence is targeted at the group controlling the village, with the aim to defeat them.

To measure taxes, I collected both whether a tax was levied on a given activity, and the level of the tax. There was always consensus on whether an expropriation was a tax. Local populations are familiar with distinct types of expropriation by various actors. While a tax is always anticipated, other types of expropriation (*pillages*, or arbitrary expropriations, known as “tracasseries”) are not.

Having collected the geographic coordinates during the survey, I can link my dataset to geographical shapefiles I obtained from the Référentiel Géographique Commun.23 This source contains the map of the road network of the DRC, all airports (including small landing lanes), the location of forests, rivers, lakes, and the regional capitals. I compute the shortest distance of each village to the road, the lake, the forest, the regional capitals, Rwanda, and the closest airport. In the analysis, I use a dummy variable indicating whether the distance to the closest airport is above the 50% percentile in the sample. The results are unchanged when I use the continuous measure. The next section describes the strategies implemented to address measurement error and potential reporting biases at the data collection phase.

---

4.3 Design based strategies to minimize measurement error

Recall data and mis-reporting of sensitive information may lead to large measurement error (de Nicola and Gine, 2012). To address this at the design phase, I used three strategies. First, I designed the village protocols so that surveyors received multiple signals on the village history from private conversations during the 7 days of implementation. Surveyors compiled the information collected in household surveys with the data assembled by the village specialists. This strategy was highly effective as an averaging method. It also helped surveyors correct information provided by the village specialists during the final meeting. Indeed, households were overwhelmingly willing to reveal armed groups’ data in private, and there was consensus that this was not problematic. To provide additional confidence on the survey measures, Section 7.1 compares the data to other data sources, as well as well-known political events. Second, the teams implemented an exhaustive set of time cues, based on information gathered in the survey pilot, and on well known regional and national political events. Time cues allowed respondents and surveyors to assign with certainty events to years, based on common regional contextual knowledge. Third, based on the pilot villages, survey questions focused on transitions and events easy to memorize. For instance, there are two types of bags that carry the heavy minerals (50kg and 75kg). To measure output taxes, surveyors obtained the fee paid on one bag and bag size used. I then compute the tax per kilogram based on these variables. To measure taxes on labor, surveyors obtained the daily fee paid at the entry of the mine to obtain the right to work. These fees are stable. Crucially, because mining taxes are the heavier daily expense of miners, they were straightforward to collect. The responses of households and village specialists were highly correlated at time of the interviews.

4.4 Descriptive statistics

Through the period, stationary bandits settled in individual villages provide security, administer justice, or establish an administration in many of the sampled villages. Figure 5 shows armed groups’ presence in the sample villages over this period. There are 18 armed organizations in the sample, although most of the variation is concentrated in four groups: the Congolese Army (40% of control year*villages and 5% of attacks), Mayi-Mayi’s (20% of control year*villages and 21% of attacks), RCD-Goma (15% of control year*villages and 11% of attacks), and FDLR (12% of control year*villages and 45% of attacks). While the RCD and the Mayi-Mayi’s were the main actors during the Second Congo war, the Congolese Army and the FDLR are 85% of the occupation year*villages after 2006. Figure 6 shows the distribution of length stationary bandits’ episodes in the period for which I have data. The median length of episodes with each stationary bandit is 4 years, exactly as the median unprotected episode. Table 1 shows summary statistics.

---

24 Armed groups were absent in most villages. The Congolese army already protected a large number of the sampled villages. In the Territoire of Shabunda, certain villages were under the control of the Raia-Mutombokis at time of the interview. The Raia-Mutombokis did not consider sensitive the information collected by this survey. This armed group had existed for less than one year. Surveyors collected historical data on events preceding the Raias and about the Raias without problems. Surveyors were confident that they achieved the same degree of quality as in other villages, including for data on the present years.
The effect of prices on the returns to form of states

This section examines the effect of the coltan and gold price shocks on armed groups activity. I first present the econometric specification and discuss how I address potential threats to identification. I then turn to the results of the baseline specification as well as their robustness to alternative specifications. I examine the properties of armed groups’ taxation in Section 6 in order to confirm the main theoretical channel.

5.1 The effect of prices on the returns to form of states: Econometric specification

The main causal effect of interest is the effect of mineral demand on armed groups’ conquest and emergence of new village stationary bandits. To measure the rise in the value of acquiring a violence monopoly, I use the following variables: conquest operations aimed at capturing the village; emergence of new stationary bandits; investment in deterrence (acquisition of weapons). However, identifying demand shifters poses a challenge. First, local prices are endogenous and might reflect changes in supply. For instance, by exercising local monopoly power or depressing supply, armed groups might inflate local prices. Second, local prices are retrospective, and despite efforts to increase the quality of data collection, they might contain large measurement error; this may be systematically different across minerals or periods. I use the world prices instead. The main specification is an OLS regression on the world prices interacted with dummies indicating the (constant) mineral endowments at the village-level, with village and year fixed effects. While constant unobserved heterogeneity and common year effects are absorbed in the fixed effects, imbalance between coltan villages and gold villages could be associated with differential time trends. To reassure that coltan and gold villages are comparable on observable characteristics, Table 2 shows the balance on observables between coltan and gold villages. Of 11 outcomes considered, there is only imbalance in one (distance to closest bridge in 2010), although this variable is post-treatment. In Section 5.2, I describe alternative methods I have used to analyze rare events data. Equation 2 presents the baseline specification:

\[ Y_{it} = \beta_t + \alpha_i + \gamma C_i P_{ct}^{world} + X_{it}' \beta + \varepsilon_{it} \]  

(2)

where \( \alpha_i \) are village fixed effects, \( \beta_t \) are the year fixed effects, and \( X_{it} \) is a vector of village level time varying controls. \( C_i \) is a dummy for whether village is endowed with coltan, and is constant over time. \( P_{ct}^{world} \) is the world price of coltan at year t. Regressors \( P_{ct}^{world} \) and \( C_i \) are absorbed respectively by the year and village fixed effects. I do not use time-varying controls, since these are potentially endogenous to treatment and could lead to bias. However, I address a number

\[ \text{The results obtained from instrumenting local prices with world prices are identical.} \]

\[ \text{This is the “bad control” problem in Angrist and Pischke (2009).} \]
of omitted variable problems by estimating the time-varying effects of time invariant covariates, which I describe in this section. To account for over-rejection stemming from autocorrelation of errors and prices, I cluster the standard errors in all regressions at the level of the village.27

A test of prediction \(P_1\) is \(\gamma_c \leq 0\) against the alternative \(\gamma_c > 0\). In what follows, I refer to \(\gamma_c\) as the main coefficient. To test prediction \(P_1'\), I include a dummy indicating whether the distance to the closest airport is above the sample median. I choose this strategy for ease of interpretation and to avoid making linearity assumptions if I instead used the continuous distance variable. The results are identical when I use the distance in kilometers. Equation 3 presents the specification I use to test prediction \(P_1'\):

\[
Y_{it} = \tilde{\beta}_i + \tilde{\alpha}_i + \tilde{\gamma}_c C_i P_{ct}^{world} + \tilde{\gamma}_{ca} C_i P_{ct}^{world} d_{i} + \tilde{\gamma}_{pa} P_{ct}^{world} d_{i} + X_{it}' \tilde{\beta} + \tilde{\varepsilon}_{it} \tag{3}
\]

where \(\tilde{\alpha}_i\) are village fixed effects, \(\tilde{\beta}_i\) are the year fixed effects. The term \(d_{i}^a\) is a dummy variable indicating whether the time invariant distance to the closest airport is above the median closest distance to an airport in the sample. Prediction \(P_1'\) therefore implies that \(\tilde{\gamma}_c > 0\) and \(\tilde{\gamma}_{ca} < 0\). Finally, when I include the main interaction for gold, \(\gamma_g G_i P_{gt}^{world}\), a test of prediction \(P_2\) is \(\gamma_c \leq \gamma_g\) against the alternative \(\gamma_c > \gamma_g\).

In what follows, I discuss the threats to identification of the main effect (\(\gamma_c > 0\)).

One may worry that the coltan shock could be endogenous. This is implausible. The source of the coltan shock is very well documented. A demand shock in the coltan processing industry in response to innovations in video-games generated a rush for coltan in DRC. It is therefore very unlikely that the world price armed groups actions had any causal effect on the world price shock.28

A plausible worry is that, given it only takes place in one year, the timing of the coltan shock may coincide with events that occur systematically in coltan villages for reasons unrelated to the price. Since the coltan shock occurred during an episode of war, it is possible that the difference in conquests between coltan and gold villages is larger during the war than on average. This will be the case if geographic characteristics interact with coltan endowments and are related to the war but unrelated to the price shock. For instance, the presence of the Congolese Army in the

---

27Bertrand, Duflo, and Mullainathan (2004) demonstrate that autocorrelations in the treatment and outcome lead to underestimation of the standard errors in studies based on differences in differences.

28As a validity check on the relevance of the coltan shock, I collected economic data in the households surveys. The retrospective households’ economic data reveals large economic effects of the mineral price shock. Figure 20 shows the number of marriages per village, yearly as collected in the village survey. Since marriage in the survey area requires the payment of a Bride Price, marriage is a normal good. The coltan shock drastically increased the number of celebrated marriages and led to large reallocation of labor to the mining sector. Second, nighttime satellite lights data confirm the evidence of the coltan shock. Figure 19 shows that a new bright town emerges in the year 2000 near Goma, the capital of the coltan province. This lighting is absent prior to 2000, and vanishes progressively in the following years. Henderson, Storeygard, and Weil (2011) introduced the nighttime data as a proxy for GDP. While not reported, this lighting is reflected in the provinces average stable lights. In addition, I computed zonal statistics of stable lights in each Province of the DRC and compared the change over time between coltan provinces and non-coltan provinces. The results show an increase in stable lights in 2000 in coltan Provinces and Rwanda, but remain constant in the remaining provinces and Uganda. This is consistent with historical accounts of the two countries role in the mineral trade and sponsoring of armed groups in DRC.
second period might act as a deterrent on armed groups conquests in coltan villages. To address this, in regressions on the whole period, I include controls in the specifications for presence of the Congolese army. To measure the presence of the Congolese army, I use a dummy indicating whether the army controls a given village in a given year. In addition, I also include the proportion of neighboring villages that are under the control of the Congolese army in a given year. I operationalize “neighboring villages” by estimating the average number of sample villages in the same administrative division that are under the control of the Congolese army, replicating this procedure for all levels of the administration. Since the results are identical with all strategies, I include only the village-level variable. However, the presence of the Congolese army is potentially a bad control (Angrist and Pischke, 2009). For robustness of the main estimate, I therefore run the baseline regression restricting the years to the Second Congo War (1998-2004), when the Congolese army is absent and the identity of the main armed groups is stable. Furthermore, since the coltan shock occurred during an episode of war, the sequence of conquest could reflect omitted variables that are correlated with mineral endowments. In particular, even when the analysis is restricted to the war years, it could be that coltan villages are more remote, and that armed groups first occupied villages near the road, followed by coltan villages. In that case, the timing of conquest of coltan villages could coincide with the timing of the coltan shock. This is unlikely. As Table 2 shows, coltan and gold villages are not distinguishable on any observable geographic characteristics, except for the distance to bridges in the year 2010. To also account for geographic omitted variables related to the sequence of conquest, I replicate the results including as controls year dummies interacted with distance to the road, distance to airports, bridges, parks, lakes, and main trading towns. These controls capture the time-varying impacts of time-invariant geographical variables that may predict the timing of conquest. I only report the results based on the interaction with distance to the road, since results are unchanged when I use the other variables instead. In what follows, I discuss threats to the interpretation that $\gamma_c - \gamma_g$ is caused by heterogeneous tax base elasticities.

The coltan shock and the gold shocks are of different magnitudes, and take place at different times. This might invalidate interpretations of $\gamma_c > \gamma_g$ in terms of tax elasticities. Shocks of different magnitudes might generate very different armed groups’ responses if the armed groups’ cost functions are concave. For instance, if there is a fixed cost of engaging in conquest, a larger price shock might result in stronger outcomes for reasons unrelated to the physical properties of production of the mineral, but only related to the magnitude of the shock. Furthermore, the timing of the shocks may be correlated with omitted variables. For instance, following the Second Congo war in 2003, the Congolese Army progressively regains control. If the Congolese Army acts as a deterrent, even when the difference in differences identifying assumptions for each the separate effects are valid, the estimated value of $\gamma_c - \gamma_g$ would be biased. To address this, I follow the same two strategies that I introduced to estimate the main coltan effect: I control flexibly for presence of the Congolese army, and I restrict the analysis for the Second Congo War years.

Another possible threat to the interpretation that $\gamma_c - \gamma_g$ is caused by heterogeneous tax base elasticities is that, even if coltan and gold villages are distributed “randomly” in the space, populations might not. If populations self select into gold or coltan villages, the patterns of sorting by
individual characteristics might be associated with different individual treatment effects of price shocks. In particular, it could be that because of the features of gold production, miners and armed groups specializing in gold have different endowments and might react differently to price shocks. This a weaker concern for the effects on conquest operations, since the decision to engage in conquest is made by bandits outside the village. In what follows, I discuss bias arising in the estimation of the standard errors.

Finally, since the econometric strategy exploits price changes across multiple years, standard errors may be underestimated. This may lead to over-rejection. If neighboring villages are subject to common shocks, an error structure with positive intra-cluster correlation within village clusters could drive differences in outcomes for reasons unrelated to the timing of the shocks (Moulton, 1986). It is unlikely that spatial clustering leads to bias in OLS standard errors, because the distribution mineral endowments is balanced within most administrative divisions, as shown in Figure 3. For robustness, I replicated all regressions clustering the standard errors at the level of the year interacted with the lowest level of the administration above the village (Groupement). To account for autocorrelation of village errors over time, all regressions are presented with clustering of standard errors at the village level. To go a step further, I use randomization inference (Gerber and Green, 2012). This allows me to account for arbitrary error structures underlying the main coltan shock. To implement randomization inference, I simulate a large number of fake allocations of the world prices to years and recompute for each price allocation the corresponding coefficient from the OLS estimation. This generates a distribution of coefficients estimated through OLS. Under the null hypothesis that the coltan price has no effect, and ignoring violations of the strict exogeneity assumption, this is the true distribution of the estimated coefficient in the sample. This also allows me to avoid making assumptions on the distribution of errors. If there is a problem of over-rejection because of the error structure, the real coefficient should lie within the 95% of simulated coefficients around the mean. I also use randomization inference to estimate the effects of the distance to airports, given the inherently spatial structure of the distance problem. I next turn to the results.

5.2 The effect of prices on the returns to form of states: Results

This section implements this econometric strategy and takes the predictions of the effect of price shocks to the data. I first describe the results, and in a second part, I show that the results are robust to alternative econometric strategies.

---

29 Groupement is the lowest administrative division above the village. It generally comprises no more than twenty villages. There are 31 Groupements in the sample, which generate more than 300 clusters for regressions including all years, and 180 clusters for regressions including only the Second Congo War. In addition, I replicated the regressions including coltan*year clustering of standard errors to account for any correlation structure among villages with the same coltan endowment that might not be captured in spatial shocks. This could be the case, for instance, if coltan villages have a common shock structure through economic networks that are not tied to geography.
5.2.1 Baseline results

The theoretical model considers how changes in the prices of outputs in individual villages affect the choices of armed groups in these villages. In particular, when the value of the village taxable production increases, armed groups will attempt to acquire the village. If the village is unprotected, they might simply march in the village, but if an armed group already controls the village, armed groups may attempt to conquer the village. If the conquest is successful, they will emerge as a new stationary bandit in the village, and collect taxes. I use these outcomes to test the main theoretical predictions: (P1) the coltan shock leads armed groups to attempt to acquire the coltan villages; (P1') this effect must exclusively take place near airports; (P2) the effect of the gold price shock is lower than the effect of the coltan shock. 30 I start with the analysis of territorial conquests.

Figure 8 shows the proportion of mines in the sample in which armed groups attacked, for gold villages (left) and for coltan villages (right). While the proportion of gold villages attacked stays around 15% between 1999 and 2000, the proportion of coltan mines attacked increases from 13% to 40%. I next show that these attacks are motivated by the conquest of territory, not pillage. Figure 9 is identical to Figure 8, but uses conquest operations instead, which I observe for the village hosting the mine. The spike in conquest attempts in the year 2000 in coltan villages suggests that the coltan shock led to an increase in territorial conquests.

Table 3 presents the results on conquests from the OLS regression. Starting with prediction P1 and P1', the results confirm that the coltan shock led to an increase in conquest attempts in coltan villages near airports. Columns (1)-(3) present the results based on the whole period, while columns (4)-(9) restrict the years to the Second Congo War (1998-2003). Column (1) presents the results from the baseline specification. As expected from prediction P1, an increase in the coltan price leads to an increase in conquest attempts on coltan villages. Column (2) adds an interaction of the main explanatory variable (Coltan X Coltan Price) with a dummy indicating whether the village is above the median distance to the closest airport in the sample. As expected from prediction P', the effect of the coltan price on conquests is concentrated near airports. Column (3) controls for the presence of the Congolese army. One might indeed worry that since columns (1) and (2) use the full period, unobservables related to the timing of the coltan shock may be associated with armed groups' activity for reasons unrelated to the price shock. If, in the period following the Second Congo War, the Congolese army acts as a deterrent on armed groups especially in coltan villages, estimating the effect of the coltan shock considering all years may be biased. While the coefficient on the Congolese army dummy is significant, including it leaves the main result unaffected. The results are identical when instead of the presence of the Congolese army I use three alternative proxies for presence of the Congolese army, described in Section 5.1.

Columns (4)-(9) present the results of the baseline specification when the sample is restricted to the years of the Second Congo War (1998-2003). The Congolese army is indeed potentially a “bad

30 A possible concern is the airports may be constructed at a low cost. The qualitative fieldwork suggests that mining areas were often located in remote forests under the control of numerous adverse armed groups. This made building new landing lanes impossible. It is not implausible, however, that had the coltan shock lasted longer, the armed groups might have engaged in creating new airports.
control” (Angrist and Pischke, 2009). Restricting the sample to the years of the Second Congo War allows me to eliminate omitted variables correlated with the change in environment associated with the post-war period. Comparing column (4) to the baseline specification in column (1), the main coefficient is larger in column (4). Column (5) includes the interaction with the distance to airports dummy. The coefficient on the main interaction (Coltan X Coltan Price) is larger, while its interaction with the distance to airports dummy is negative. This confirms that the effect of the coltan price is concentrated near airports. Column (6) controls for distance to the road, interacted with year dummies, to account for flexible time-varying effects of distance to the road. I do this because the sequence of conquests may be associated with geographical characteristics of coltan and gold villages unrelated to the coltan price (remotedness). The inclusion of this control leaves the coefficient on the coltan and coltan price interaction unaffected. To account for year region-specific shocks, column (7) includes region-year fixed effects. One could indeed worry that the best coltan villages are concentrated in one region, where most conquests take place in the year 2000 for reasons unrelated to the shock. Including region*year fixed effects leaves the main coefficient unaffected.

Columns (8) and (9) restrict the sample to 1999 and 2000. I exclude all other years to prevent the serial correlation across years to lead to under-estimation the standard errors. If we use information from many years around the coltan shock, we may indeed understate the standard errors on the coefficient estimates (Bertrand, Duflo, and Mullainathan, 2004). When I restrict the years to 1999-2000, the “treatment” spans only one period. I therefore do not need to cluster the standard errors at the village level. The coefficient is still significant, and as column (9) shows, the main effect is concentrated near airports. My baseline estimate suggests that the coltan shock increased the attempted conquest attacks by 20%.

Turning to prediction $P_2$, an increase in the price of coltan should have a larger effect than an equivalent value per unit increase in the price of gold. The point estimates provide support for this prediction across columns. The gold price is negatively related to conquest operations in gold villages in the whole period (columns 1-3). This negative relationship turns zero when I consider only the years of the Second Congo War (columns 4-9), or when I include village linear time trends (yielding a p-value of 0.68), suggesting it is likely due to unobservables. Confirming these results, a t-test of the null hypothesis that the coefficient on coltan is smaller or equal than the coefficient on gold in the baseline specification rejects the null hypothesis with a p-value of 0.03. However, while the discrepancy in coefficients is consistent with the theoretical results, it is no evidence that the underlying mechanism is the inability to tax gold output, since this result obtains even when gold output can be taxed (because of its smaller volume productivity). I then implement a test rescaling the estimated coefficients obtained from the baseline specification as derived in Section 3.3, and find evidence that the inability to tax gold output alone can explain the magnitude of the coefficients. I indeed reject the null Hypothesis with p-value 0. In sum, the results on conquest attempts confirm predictions $P_1$-$P_2$.

Table 5 runs the analysis of Table 3 looking instead at whether a new stationary bandit acquires the monopoly of violence in the village. The “entry” variable I use here is thus the union of

---

31 Results with time trends are unreported.
emergence of stationary bandits in areas where there was previously no stationary bandit, and the emergence of a new stationary bandit in areas where there was already one stationary bandit. The results, across specifications, are identical. Thus, the price shock effectively leads to the entry of new stationary bandits as a result of conquests.

I now turn to the results on pillages. Following predictions P1-P2, the price shock should not increase pillages, and if at all, it should decrease pillages by the strong armed group. Table 4 runs the analysis of Table 3, for pillages instead of conquests. Across all specifications, the estimates show that pillage operations are unaffected by the coltan shock. Consistent with the model, an increase in the price of coltan increases fighting for control of coltan villages, not pillaging.

Finally, Tables 6 and 7 present the same specifications for the complete set of outcome variables that characterize the transitions of individual villages in and out of stationary bandits occupation. I focus on the following variables: presence of a stationary bandit; transitions of villages from stationary bandit to no stationary bandit; presence of taxation by armed groups in the village; acquisition of firearms by armed groups (which captures investment in deterrence of invaders). Since the coltan shock and the gold shock take place in different years, I separate the estimations for each shock. For the coltan shock, I focus on 1999 and 2000, which is the most conservative set of years. For the gold shock, I use the years 2001-2003. The results are invariant to the choice of years, and adding later years to the gold specification improves my results. Across Tables 6 and 7, columns (1), (3), (5), (7), report the results considering the baseline specification, while columns (2), (4), (6), (8), include the interaction with the distance to airport dummy. The results are consistent with the interpretation of the findings on conquest attempts: the coltan price shock increases the value of acquiring monopolies of violence, especially near airports. When the village is empty, it is more likely that a stationary bandit emerges, while if the village already has a stationary bandit (which was the case for 55% of coltan villages already by 1999), armed groups attempt a territorial conquest. These attempts are sometimes successful, in which case armed groups that dominate the local resistance acquire the monopoly of violence, collect taxes, consistent with historical accounts of state formation in medieval Europe. Table 7 shows that this effect is absent in gold villages. Figures 10 and 11 present a summary of results in tree form. They separate outcomes by whether the village already had a stationary bandit and otherwise. I next consider alternative specifications.

5.2.2 Robustness: alternative specifications

In addition, I implement three strategies to increase confidence in the main OLS estimates. First, I implement conditional differences-in-differences matching, introduced by Heckman, Ichimura, and Todd (1998). As matching variables, I use distance to airports and distance to

---

While investment in deterrence should unambiguously increase in the presence of strategic interactions, it may reflect an income effect. As the price rises, groups in coltan villages may have more means to acquire weaponse. For that reason, acquisition of firearms is not unambiguously indicating that the value of territorial control has increased.

This method is similar in spirit to case control methods, presented in Goldstone, Bates, Epstein, Gurr, Lustik, Marshall, Ulfelder, and Woodward (2010), where I select observations based on whether I observe coltan endowment and then select matches.
roads, as well as upper level administrative divisions (Territoire). Since airport areas might be more urbanized, and urbanization might affect the behavior of armed groups, I also consider distance to roads. Finally, to control for constant unobserved heterogeneity across regions (Territoires), I also match on Territoire. This procedure is thus equivalent to conditional differences-in-differences within calipers defined by administrative divisions (Cochran and Rubin, 1973). Figure 12 shows the yearly coefficients estimated on the matched sample, based on the main outcome variables: conquest attempts, pillages, entry of new stationary bandits, and acquisition of firearms. The results are identical.

Second, I account for the observation that conquest attempts are rare events. In that case, OLS estimation could lead to biases in the estimated probabilities and standard errors (King and Zeng, 2001). I therefore run a logistic version of the baseline specification, that corrects for small sample and rare events in order to generate approximately unbiased and lower-variance estimates, introduced in King and Zeng (2001). The results are unchanged.

Third I account for the spatial problem in regressions that include the distance to airports. Neighboring villages have similar distances to the closest airport, and they are also likely to be subject to common shocks stemming from their economic and social proximity. The problem is not limited to villages in same clusters. Villages close to an airport, even if they are close to different airports, may be subject to common shocks, if they are better integrated in the regional and world economy. This error group structure can lead to drastic under-estimation of standard errors in the OLS framework. One way to address this is to cluster the standard errors at the level of “Groupements” * years, which could account for a grouped error structure if errors are subject to common year shocks within “Groupements”. As announced in Section 5.1, I implemented this for all regressions, and the results remain unchanged. However, the particular error structure is unknown, and therefore the correct clustering is also unknown. Alternatively, one could use Conley standard errors, which account for an error structure of AR(1) type (Conley, 1999). However, estimation of a spatial AR(1) structure on the errors will not capture arbitrary correlations. This problem is akin to imposing a serial AR(1) in the errors for panel regressions (Bertrand, Duflo, and Mullainathan, 2004). To exploit the spatial structure inherent in the data, I instead use randomization inference, which I take from the statistics literature. While rarely used in economics, it allows to exploit the structure of the data without making assumptions on the data generating process. I generate 5,000 times 40 imaginary airports’ locations in the space around the survey villages. For each of these simulated airport locations, I run the OLS regression, where I interact the distance to airport dummy with the main regressor (Coltan X Coltan Price). I then estimate the coefficient on the triple interaction for each simulation, and plot the resulting distribution of coefficients in Figure 13. Since these airports are not real, and thus have no effect in expectation, the distribution of the coefficients is centered around 0. If a spatial structure leads to over-rejection, the distribution of coefficients estimated through simulations should have a relatively high variance. My coefficient estimated with real airports would lie within the 95% of the data around the mean. On the contrary, I find that only 9% of the coefficients obtained through simulations are larger in absolute value than the coefficient estimated based on the real airports, a 0.09 p-value of a two sided test.
of statistical significance. The main result, therefore, remains when I estimate the standard errors using randomization inference.\textsuperscript{34}

This section has shown that the coltan shock leads armed groups to acquire control of coltan villages near airports and collect taxes. The later shock to the price of gold, which can easily be hidden from taxes, does not. The next section examines the activity of armed groups in the village where they are stationary bandits.

6 Understanding the economic incentives of \textit{stationary} bandits

To strengthen the main results, this section examines the relationship between armed groups and the economy in the villages they control. I establish three sets of results. i) If the two main results are explained by the returns to tax, the taxes raised by \textit{stationary} bandits should reflect that gold is harder to tax. Furthermore, the \textit{stationary} bandit’s tax should minimize distortions. ii) If the two main results are explained by the returns to tax, then a decrease in the returns to tax stemming for a decrease in the time horizon should revert the main result. iii) If the two main results are explained by the returns to tax, then armed groups should use violence promote economic activity.

6.1 Taxation by \textit{stationary} bandits

This section examines the taxation behavior of the stationary bandits, and tests the predictions derived from optimal taxation.

Armed groups use taxes on various observable economic activities. Table 8 shows the proportion of villages in the sample in which armed groups collect taxes on the different economic activities, separating villages that are exclusively endowed with coltan from villages that are exclusively endowed with gold. Armed groups raise taxes on mineral output, mining labor, agricultural production (taxes on food sales at the market), transit of persons (roadblock at village exit), and on activity of the local mill. Table 8 also shows that armed groups use poll taxes, consistent with the predictions on poll taxes. Table 9 shows the OLS regressions of the choice of taxes on prices interacted with mineral endowments. Positive shocks to both gold and coltan lead to the emergence of taxes on mineral activity.

The results confirm that gold cannot be taxed, the main channel leading to prediction \textit{P2}. Table 8 shows that in gold mines, \( \tau = 0 \) and \( t > 0 \), consistent with assumptions G1 and G2 that emphasize the ease of hiding gold. Table 10 shows that the price shocks increase the mining sector tax rates, as expected if the price shocks reduce labor supply elasticities.

I also find evidence that the \textit{stationary} bandit’s taxes are consistent with the fiscal externality channel. In order to estimate \( \text{cov} \left( \tau_i, \sum_{i \neq l}^{N} \tau_i p_i Y_i \right) \), I computed the day equivalent tax in dollars in various sectors. For the mining sector, I took use taxes at the mine of the village, while for the other sectors I also implemented this procedure simulating allocations of mineral endowments to villages as well as price to years. Figure 13 shows that the estimated p-value with randomization inference is 0.065.

\textsuperscript{34}
agricultural sector I take taxes at the market on cassava sales. I then computed a daily sector i tax index for each sector as follows:

\[ I_i = \left( \hat{Y}_i \tau_i + t_i \right) p_i \]

where \( \hat{Y}_i \) is the estimated average volume product per worker for one day of work, \( \tau_i \) is sector i tax on output, \( t_i \) is sector i tax levied on one day of work, and \( p_i \) is the local output price in sector i.\textsuperscript{35}

Since the daily tax indexes have a large mass of observations at 0, I use a tobit regression of daily tax index in the relevant mining sector, on the daily tax index in agriculture, clustering standard errors at the village level. Table 11 presents the result. There is a positive association between the mining tax index and the estimated tax revenues from agriculture, consistent with the prediction from the model.

While I do not observe regulation of village entries and imposition of local markets in the dataset, the qualitative evidence drafted by the surveyors demonstrates that settled armed groups use both strategies on a widespread basis. Furthermore, at mines, armed groups make efforts to increase the costs of hiding mineral output. They use violence in order to racket mineral output when civilians are hiding it, and confiscate clothes in order to inspect them.

Finally, the mix of input and output taxes is consistent with the theoretical predictions under the standard assumptions on risk aversion and imperfect monitoring. Figures 14 and 15 show that armed groups provide a mix of input and output taxes around coltan mines. This elaborate contract is consistent with optimal taxation if both the bandits and the producers are risk averse and their incomes are tied to mineral extraction. Both assumptions are reasonable and match the anecdotal and qualitative evidence on the mining sector.\textsuperscript{36}

Broadly, this section suggests that armed groups’ taxation is consistent with the main channel of the model (gold output is not taxed), as well as with the predictions associated with optimal tax. This has an important implication: settled armed groups internalize the behavioral responses of the population and intend to minimize distortions, confirming that they behave as Olson (1993)’s stationary bandits, instead as short-sighted pillagers. This will depend, however, of how long they expect to stay stationary bandits. The next section examines the effect of the time horizon using an external natural experiment.

6.2 Stationary bandits and the time horizon: effects of the Kimia II operation

Section 5 showed that in response to mineral demands, armed groups fight to obtain taxation monopolies where they then raise taxes consistent with optimal taxation. However, if armed groups internalize the future effects of current extraction, as the time horizon shrinks, the value of the monopoly of violence also shrinks, which should revert the main result (P1). This section shows

\textsuperscript{35}Since I do not observe taxes on output for gold and since the tax on cassava is charged on one day of sales at the market, \( \tau_i = 0 \) in gold and agriculture. I therefore only need to estimate production for heavy minerals, coltan and cassiterite. Based on the average production reported in all year-village observations, I imputed 20 kilograms both for cassiterite and coltan.

\textsuperscript{36}With non-linear tax instruments, bandits could set tax rates that depend on production. It is however unlikely that this type of contract is implementable, especially with the technology at hand.
that stationary bandits indeed revert to pillage of assets when their time horizon shrinks.

I exploit the timing and targeting of a military strategy that cut the public finance of the Forces De Libération du Rwanda (FDLR). In 2009 and 2010, the Congolese Army, with the help of United Nations forces and the Rwandan army, targeted villages under FDLR control in order to weaken their link with the tax base. The Congolese army scaled up the operation in 2010, which took the name of Amani Leo. Henceforth, I refer to both as the Kimia II operations, consistent with the villagers’ parlance.\(^{37}\) On the one hand, the threat posed by Kimia II suddenly reduced the time horizon of the FDLR. On the other hand, consistent with historical accounts of the formation of fiscal capacity during war (Tilly, 1985), the FDLR might have had incentives to increase tax collection in order to secure their territorial control. The second effect is implausible. The FDLR was very weak and had no hope of resisting the operation. It is natural to interpret the operation as a shock to the time horizon. The econometric specification is:

\[
Y_{it} = \alpha_i + \beta_t + aPOST_t \ast FDLR_{2008}^i + \varepsilon_{it}
\]

where \(i\) indicates the village and \(t\) indicates the year. \(Y_{it}\) are village level pillages by the FDLR. \(\alpha_i\) are village fixed effects, \(\beta_t\) are year fixed effects, \(POST_t\) is a dummy taking the value 1 for all observations after 2008 and 0 otherwise, and \(FDLR_{2008}^i\) a dummy indicating whether a village was under the control of the FDLR in 2008. The coefficient of interest, \(a\) captures the increase in pillages since Kimia II operations in FDLR villages (henceforth, treated villages), compared to the increase in pillages since Kimia II operations in non-FDLR villages. A test of the prediction is thus \(H_0: a \leq 0\) against the alternative \(H_1: a > 0\). Figure 16 presents graphical evidence on the intervention. Following the Kimia II intervention, FDLR taxation drops from 100% to 20% in the former FDLR villages within 2 years.

Figure 16 also presents the results. While the number of FDLR pillaging operations remains around 10% of villages in the control group, it rises from 0% to 30% in the treated group (the former FDLR villages). Table 12 provides Difference in Differences estimates. Pillaging operations by the FDLR increase drastically against their formerly controlled villages. These operations take place during the night and are brief. They can therefore be implemented despite the presence of the Congolese Army in neighboring villages, while stable taxation cannot. It is well documented that during the Kimia II operations, the FDLR relocated in the neighboring forests, and increased their pillages on the villages they used to control in order to guarantee access to resources. These results are also strongly supported by the qualitative reports elaborated by the surveyors. The villagers narrations collected in the surveyors’ qualitative reports indicate that the increase in

\(^{37}\)In South Kivu, Kimia II operations also resulted in some important successes as FARDC gained control of much of Kalehe, Kabare and Shabunda territories after years of domination by FDLR. On 28 July, FARDC succeeded in dislodging FDLR from Kashindaba, its main headquarters in South Kivu. In August, FARDC operations challenged FDLR strongholds in Walungu, Mwenga and Sange and around Uvira. On 20 August, FARDC successfully concluded operations to clear FDLR presence from the Kahuzi-Biega National Park as well as Tehivanga and Nindja areas in Kabare territory\(^{37}\). Following Kimia II, the Congolese army continued its implementation under the label of Amani Leo. “Amani Leo also aims to regain control of mining sites from the FDLR and combat illegal trafficking in minerals and other natural resources.” Source: http://monusco.unmissions.org.
pillages was caused by the arrival of the Kimia II operations, which shortened the hopes of the FDLR to perpetuate taxation. Before fleeing, the FDLR pillaged.\textsuperscript{38} This result suggests that the time horizon, by affecting the expected revenues from taxation, is a determinant of the behavior of stationary bandits.

6.3 Economic impact of stationary bandits: effects of a peace agreement

Armed groups might just be fighting over assets that increased in value. However, if they internalize the value of economic activity, they should provide protection and promote growth.

This section examines empirically the economic effects of stationary bandits in the village they occupy. If armed groups have a comparative advantage in protection, they will have an interest in protecting local production in order to tax it when they are stationary bandits. If property rights are sufficiently insecure in the absence of a stationary bandits, his protection will have a positive effect on output. In addition to economic effects, I present the results using indicators of the population welfare. Control by a stationary bandit has ambiguous distributional implications: while they provide protection, they also maximize tax revenues.

Since the price shocks lead to stationary bandits, I could have instrumented presence of a stationary bandit with the coltan shock in order to estimate the effect of stationary bandits on the local economy. However, such a strategy will obviously violate the exclusion restriction since the price is likely to be associated with economic activity independently of armed groups’ behavior. Instead, in order to estimate the impact of stationary bandits on the local economy, I exploit the timing of a Peace agreement signed in 2003 that was intended to affect all villages controlled by the Rassemblement Congolais pour la Démocratie equally. This agreement led the RCD to abandon the villages it used to control. The withdraw of the RCD resulted in a security vacuum. The villages that were under the control of the RCD in 2003 are sparsely distributed across the survey area. By 2005, the RCD had abandoned all villages. The two stage least squares identification strategy is:

\[
Y_{it} = \alpha_i + \beta_t + aRCD_{it} + \varepsilon_{it} \\
RCD_{it} = \eta_i + \gamma_t + bPOST_i \ast RCD^{2002}_{i} + u_{it}
\]

where \(i\) indicates the village and \(t\) indicates the year. \(POST_t\) is a dummy taking the value 1 for all observations after 2002 and 0 otherwise, and \(RCD^{2002}_i\) a dummy indicating whether a village was under the control of the RCD in 2002. \(Y_{it}\) are village level outcomes, in particular, dummies indicating mining activity and recorded output; \(\alpha_i\) are village fixed effects, and \(\beta_t\) are year fixed effects. Since the RCD location is potentially endogenous, I use the timing of the RCD withdrawal interacted with its initial location as an instrument. The coefficient of interest, \(a\) captures the changes in economic indicators in RCD villages since the Sun-City Peace agreements (henceforth, treated villages), compared to the respective changes in economic indicators in non-RCD villages,

\textsuperscript{38} External judiciary and journalistic evidence also strongly supports this finding. See http://www.icc-cpi.int/iccdocs/doc/doc1225453.pdf
under the identifying assumption that the timing of the agreement did not coincide with time trends that affected villages under the RCD control differentially. In particular, it may be that the RCD was occupying mostly coltan villages, and that the world price of coltan decreased at the same time of the agreement. This is particularly worrisome if the agreement was signed in anticipation of a decrease in prices. To account for this, I include controls for mineral endowments interacted with their prices as well as geographic controls. A test of the prediction on the impact of stationary bandits is thus $H_0: \alpha \leq 0$, estimated using the timing of the RCD withdraw as an instrument, against the alternative $H_1: \alpha > 0$, estimated using the timing of the RCD withdraw as an instrument.

Figure 17 provides a graphical representation of the first stage and the reduced form. Visually, it is apparent that the timing of the Sun-City peace agreement is associated with a mild decrease in economic activity in the villages formerly controlled by the RCD. However, this decrease may be driven by non-compliers, villages in which the RCD stayed. To address this, I instrument presence of the RCD with the timing of the Sun-City agreement interacted with the dummy indicating the presence of the RCD in 2002. Table 13 reports the corresponding IV estimates and Table 14 replicates the results with additional controls. The results suggest that the RCD presence had mild, positive effects on economic activity. The next section addresses issues measurement error concerns as well as the role of expectations, and discusses alternative explanations.

7 Discussion

In this section, I demonstrate that the outcome variables compare surprisingly well with known historical events, as well as with external data sources. I then discuss how expectations might play a central role in the main estimates, and finally address how these results fit other explanations from the literature on civil war.

7.1 Measurement error

Classical measurement error in the outcome variable might lead to noisy estimates. In addition, if the data is of low quality, arbitrary correlations stemming from systematic mistakes by enumerators might underlie arbitrary results unrelated with real effects. I first provide reassuring graphical evidence and then compare the data to the alternative data sources.

Figure 5 in section 2 plotted the survey-based measures of armed groups’ occupation on years. Since historical trends are known, this allows to verify the quality of the main variables. The changes in the data correspond precisely to commonly known historical events. Stationary bandits emerge and substitute for the state in villages in 1996 and 1997 with the AFDL rebellion, but particularly in 1998 with the RCD conquests. The data also coincides with known dates of RCD occupation: it emerged in 1998 with the second Congo War and retreats in in 2003/2004 following

39Classical measurement error in the explanatory variable leads to attenuation bias. However, to avoid this source of bias, the specifications do not use local prices of coltan.
the 2002 peace agreements of Sun City. The timing of the attacks also corresponds precisely to well known historical rebellions. Figure 21 shows that the data maps precisely to the known phases of the war: the AFDL led its rebellion in 1996 into 1997, the RCD between 1998 and 2003, and the CNDP in 2004. Finally, survey-based recall prices closely track the international world price. Overall the data on prices, armed groups, and attacks benchmarks extremely well to well-known historical events.

To verify the data quality a step further I also assigned violent events geo-coded by an external dataset (ACLED) to circles around the survey villages. The ACLED dataset contains 3,500 violence events since 1997, coded by perpetrator and type of event. As opposed to the data collected in this survey, these data are based on reports. When an event falls in circles assigned to more than one village, I allocated the event to all village circles in which it fell. I then compare this data to my source that contains the number of attacks on villages. Figure 22 shows that the ACLED dataset systematically reports less violence data around the selected villages than what the current survey. While it is possible that this difference is due to villagers over-reports of violence events, under-reporting by villagers in the survey is more likely to be a concern, due to memory loss and fear of retaliation. If data from this survey is more likely to under-report violent events, this suggests that my survey performs well, possibly improving upon the ACLED dataset for the specific locations. To cross-validate this survey data econometrically, I correlated the data to ACLED’s data in an OLS framework. Regression estimates of the correlation between ACLED and survey conquest attacks in the neighborhood of the village are positive, significant, but never larger than 10% across all distances (Table 16). In addition, attacks reported by this survey are less likely to be reported in ACLED when the villages were under RCD occupation (suggesting that the source of under-reporting in ACLED could be obstruction of information), and when the attacks were pillages (ACLED may specifically under-estimate the number pillages). The data collected in this survey also captures the well-known evolution in conflict intensity in the region. Figure 22 shows the increase in violence during the Second Congo War, especially during the coltan shock, its decrease after 2003, and its dramatic increase as a result of the Kimia II operations, which consists mostly of FDLR pillages.

7.2 Expectations

First, if villagers anticipated the coltan shock, estimates of the coltan shock may be a lower bound of the real effect. For instance, groups could have made the required investments during 1999 knowing that the price would rise. Second, following the price increase, if local populations anticipated that the shock was going to be short-lived the estimates of the effect of the shock will underestimate real effects of permanent price shocks, especially if groups’ cost functions have non-convexities. Third, following the sudden price drop in international markets, if populations expected that the

---

40 This source contains larger numbers of violent events than the other survey source because villagers could not remember the details of some attacks.

41 This result applies specifically to RCD perpetrated violent events, consistent with reporting bias.
price would rise back to its high level, the persistence of the coltan shock on institutions will be overestimated and will be led by wrong local expectations about future prices.

There is significant first-hand evidence on expectations about the coltan price level. This anecdotal evidence has three implications. First, surprise innovations on the electronics markets led to the rush to coltan. Second, the unexpected bad performance of Playstation II on the market in Christmas led the demand for coltan to vanish. It is therefore very unlikely that local populations expected a sudden drop in demand for this mineral of this magnitude. Third, following the sudden price drop in 2001, local traders expected the world demand to rise back to its peak level of 2000, and engaged in mineral hoarding. This is confirmed by the data on local prices.\footnote{United Nations Security Council (2002).}

In order to strengthen the anecdotal evidence, the survey also collected retrospective measures of expectations. The sudden increase and decrease were both unexpected. However, there is variation in expectations about the duration of the shock. Figure 18 shows the emergence of new stationary bandits. The upper quadrants show the entry of new stationary bandits. I separate villages in which villagers declared that they had expected the shock to be permanent, and villages in which villagers declared they had expected the shock to be temporary.\footnote{I have defined a temporary shock to be 5 years or less, which constitutes less than 20\% of all villages. The rest report an expected duration of more years, or reported to expect the price to remain at high level for an indefinite amount of time.} The price shock had an almost three times larger effect (entry of 50\% instead of 20\%) amongst the villages that reported to expect that the price shock would be permanent. Finally, the lower quadrant replicates the analysis by expectations on the return to high prices. Of places where villagers expected a temporary decrease, stationary bandits vacated 25\% in 2001. Of places where villagers expected no return to the high price, stationary bandits vacated progressively. This suggests that expectations may explain part of the persistence, and that had the villagers expected the prise rise to be temporary, the behavioral responses may have been dampened.

7.3 Alternative mechanisms

A large part of the conflict literature emphasizes the role of the trade-offs between production and expropriation, in particular the opportunity cost of fighting (Collier and Hoeffler, 2004, Dube and Vargas, 2013). My results suggest an additional constraint that determines the choices of armed groups: their ability to tax. The shock to the price of coltan implied a drastic increase in the economic opportunities from engaging in productive activities. Therefore, if the opportunity cost of fighting was the determinant of violence, there should have been a decrease in participation, and ultimately in violence, as found in Dube and Vargas (2013). On the contrary, I find an increase in violence, aimed at gaining territorial control. Hence, when the price of the labor intensive commodity increases, the relevant alternative for armed groups might not be to engage in productive activities in a labor intensive sector, but potentially to provide protection and extract revenue taxing output and labor. This channel is absent in the opportunity cost of fighting literature.

Second, my results complement theories of civil war that emphasize the role of the internal
organization of the group (Weinstein, 2007). As the revenues from stable extraction increase, the stronger bandit acquires control of the valuable villages. In this case, a large part of the effect (not all) is driven by the RCD. It is possible that the RCD had a better disciplined internal organization that gave them dominance over uncoordinated Mayi-Mayi groups. However, if recruitment of soldiers’ types is the only explanation for violence, it cannot explain the effects of the Kimia II. As the FDLR lost access to its mineral tax base, violence should have decreased if bad recruits had been replaced with better recruits. On the contrary, violence increased. The armed group just changed their strategy to acquire resources: they pillaged households’ assets. This result thus complements Weinstein (2007). Once established, it is the fighters’ relevant alternative revenue generating activity that determines the use of violence, which may not be production but alternative modes of expropriation.

Finally, this paper ignores strategic interactions. Future work will examine how behavior of armed groups is affected when other bandits can pillage production. If the stationary bandit does not have a sufficiently strong protection technology, this mechanism can unravel uncoordinated roving banditry. However, this is also the case in the current decision theoretic framework. This framework captures well the dominance of the RCD against other weaker groups.

8 Conclusion

While the origins of the state have attracted scholars of all social sciences, a fundamental fact in all disciplines is that prior to the state, there is no data collection capacity. There is therefore no formal statistical evidence the causes of state origins.

I provide the first econometric evidence on the causes of state formation. To observe state formation in its initial phases, I managed a team of data collectors in stateless and war-torn areas of Eastern Congo and collect a village-level panel dataset of yearly activity of armed groups. I develop a theory in which the ability to tax output plays a central role in the decision of armed organizations to acquire a “monopoly of violence” in order to tax.

Consistent with this theory, a drastic shock to the demand of coltan, a voluminous mineral easy to tax, leads armed groups to impose monopolies of violence in the proximity of coltan extraction and trade routes. A shock to gold in the immediate aftermath, does not. Since gold is easy to conceal, this is consistent with the role played by taxation in the decision to form states. As additional support to this theory, I find that armed groups, even when they are settled in gold villages, do not tax gold output.

I take advantage of two natural experiments during the time period observed to go a step further testing the theory. First, I show that a large military intervention, which reduced the time-horizon of a settled armed group reverts the main result. In response to the threat, the armed group turns to pillaging the assets of the populations they formerly protected. Second, using a political agreement that resulted in the withdrawal of a settled armed group, I show that the presence of settled armed groups is associated with higher economic activity, as expected if they are “stationary bandits”.

33
This paper’s results contribute to a large literature in economics (Acemoglu and Robinson, 2006, Acemoglu, Ticchi, and Vindigni, 2006, Bates, Greif, and Singh, 2002, Besley and Persson, 2009, Grossman, 1997, North and Weingast, 1989), but also to disciplines outside of economics. The anthropological literature on early state formation suggests that economic returns to circumscribed locations played a central role in the formation of states (Carneiro, 1970). I find that the effect of mineral demand is particularly strong in locations where output can be traded. Furthermore, since price shocks have weaker effects on output that can be easily hidden by producers, I am able to provide disaggregated evidence on the role of the elasticity of the tax. This complements historical accounts of the role of population mobility and fixed investments on state formation (Herbst, 2011, Scott, 2009).

This paper has implications for policy on armed groups, weak states, and development aid. Armed groups are a topic of uttermost interest among governments globally. This study is the first to provide econometric evidence of the relationship between conflict and mineral endowments in the DRC. While there are vivid debates and speculations around the motives of armed groups, I find that price shocks lead to armed groups violence, between groups and in order to acquire territory hosting the mineral (not to more pillages). This result is in line with other reports based on qualitative data, such UN (2002). Second, I find that interventions aimed at constraining armed groups ability to collect taxes increase violence against civilians in the short-run. Interventions that attempt to weaken armed groups finances have become dominant among policy circles. In particular, the United States issued the Dodd-Frank legislation in 2012, aimed at constraining purchases of minerals whose trade is a source of finance to armed groups. Governments interested in “cleaning” the mineral chain, thus, may need to protect civilians in the aftermath of these interventions, and provide alternative occupations to combatants who lose access to revenues from taxes as a result of these interventions.

Weak states are a major threat to global stability and a major concern for the U.K. and US administrations. First, developing the Congolese state capacity to collect taxes can be a powerful deterrent of violence against civilians by the state. The findings illustrate the role played by expected taxation revenues on government stability, reflecting existing academic scholarship on the determinants of state predation (Bates, Greif, and Singh, 2002, Besley and Persson, 2008). Second, interventions aimed at improving the Congolese state’s capacity to pay salaries to armed state officials are a promising avenue to reduce violence against civilians. Armed state officials lacking a stable source of revenue, may engage in violent expropriations in order to secure revenue. Securing stable income for state officials can be achieved by technological innovations in the salary payment chain, for instance, using a cell phone payment system.

Eastern Congo attracts a large part of the world’s development and humanitarian aid. First, this paper suggests that development aid disbursed in this area provides potentially a source of revenue to armed groups. This paper’s findings demonstrate that armed groups often settle in individual villages, and systematically raise taxes on a large number of economic activities, not only on the mineral trade, contrary to what is commonly believed. Aid that does not take into consideration the revenues it can generate to armed groups may counteract policies aimed at weakening armed
groups. Second, the findings of this paper also suggest how to design aid minimizes this risk. Aid that translates into unobservable or intangible assets, may be harder to tax. This is the case of education and health. Attention needs to be paid to the process in which these services are delivered. This also implies that the marginal returns from aid delivery will be larger if armed groups are not taxing in targeted villages, suggesting that aid delivery and interventions that eradicate armed groups are complementary. Third, aid may increase civilian exposure to risk of violence if wealth can be converted to portable assets. The findings of this study indeed suggest that pillages target specifically portable assets. This finding is consistent with the findings of Nunn and Qian (2012).

The “stationary bandits” I am able to observe are a mixed blessing. On the one hand, they weaken the authority of the state and sometimes abuse civilians. On the other hand, they engage in relationships of mutual dependency with civilians, and provide services that the state is unable to provide, in particular, protection. The results suggest they might have a positive effect on production, as long as they can derive taxation revenues from it. In addition, through the period, the state also engages in abuse of civilians. This suggests that academic and policy thinking that distinguishes stationary bandits from the state might lack a fundamental similarity between them: having comparative advantage in violence, they will choose among various forms of expropriation. These strategies might benefit the population when there is an economic surplus that can be taxed, which may in some cases lead to the formation of states.
References


Table 1: Summary statistics of main outcomes

<table>
<thead>
<tr>
<th></th>
<th>No minerals</th>
<th>Coltan</th>
<th>Gold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (sd)</td>
<td>Mean (sd)</td>
<td>Mean (sd)</td>
</tr>
<tr>
<td>Number of roadblocks where tax on output is collected</td>
<td>.09 (0.41)</td>
<td>0.00 (0.00)</td>
<td></td>
</tr>
<tr>
<td>Use of Entry fee at the mining center</td>
<td>.07 (0.41)</td>
<td>0.05 (0.41)</td>
<td></td>
</tr>
<tr>
<td>Number of external attacks at the mining center</td>
<td>0.64 (2.21)</td>
<td>0.02 (0.21)</td>
<td></td>
</tr>
<tr>
<td>Number of internal attacks at the mining center</td>
<td>0.22 (0.95)</td>
<td>0.01 (0.11)</td>
<td></td>
</tr>
<tr>
<td>Index of taxes on output</td>
<td>.03 (0.47)</td>
<td>-0.05 (0.21)</td>
<td></td>
</tr>
<tr>
<td>Use of tax on households by armed groups</td>
<td>0.36 (0.48)</td>
<td>0.38 (0.49)</td>
<td>0.51 (0.50)</td>
</tr>
<tr>
<td>Use of tax on food sales by armed groups</td>
<td>0.08 (0.27)</td>
<td>0.04 (0.19)</td>
<td>0.05 (0.21)</td>
</tr>
<tr>
<td>Index of taxes at the mining center</td>
<td>-0.74 (1.10)</td>
<td>-0.42 (1.10)</td>
<td>-1.08 (1.27)</td>
</tr>
<tr>
<td>Number of external attacks at village</td>
<td>0.18 (0.38)</td>
<td>0.13 (0.34)</td>
<td>0.22 (0.42)</td>
</tr>
<tr>
<td>Number of internal attacks at village</td>
<td>0.02 (0.14)</td>
<td>0.04 (0.20)</td>
<td>0.03 (0.17)</td>
</tr>
<tr>
<td>Armed organization stationed at the village</td>
<td>0.40 (0.49)</td>
<td>0.37 (0.48)</td>
<td>0.30 (0.46)</td>
</tr>
<tr>
<td>Armed organization living inside the village</td>
<td>0.37 (0.48)</td>
<td>0.32 (0.47)</td>
<td>0.27 (0.44)</td>
</tr>
<tr>
<td>Armed organization sharing power with the village authorities*</td>
<td>0.26 (0.44)</td>
<td>0.52 (0.50)</td>
<td>0.47 (0.50)</td>
</tr>
<tr>
<td>Armed organization administering the village</td>
<td>0.07 (0.25)</td>
<td>0.19 (0.40)</td>
<td>0.07 (0.25)</td>
</tr>
<tr>
<td>Price of cassava (C. Francs per unit of namaha)</td>
<td>455.82 (263.34)</td>
<td>552.20 (253.08)</td>
<td>570.00 (271.12)</td>
</tr>
<tr>
<td>Price of beer (C. Francs per bottle of Primus)</td>
<td>1126.80 (369.00)</td>
<td>1309.42 (505.21)</td>
<td>1150.70 (482.79)</td>
</tr>
<tr>
<td>Exchange rate (C. Francs per US Dollar)</td>
<td>834.21 (390.08)</td>
<td>859.25 (518.61)</td>
<td>804.18 (114.52)</td>
</tr>
<tr>
<td>Number of marriages per year</td>
<td>6.32 (9.52)</td>
<td>7.38 (13.37)</td>
<td>4.02 (8.51)</td>
</tr>
<tr>
<td>Size of the village (Number of households in 2012)</td>
<td>187.15 (116.39)</td>
<td>172.13 (101.38)</td>
<td>102.56 (44.07)</td>
</tr>
<tr>
<td>Number of individuals immigrating per year</td>
<td>24.4 (69.67)</td>
<td>21.37 (76.19)</td>
<td>16.19 (60.10)</td>
</tr>
<tr>
<td>Number of individuals emigrating per year</td>
<td>15.22 (49.67)</td>
<td>15.20 (59.42)</td>
<td>13.57 (52.49)</td>
</tr>
<tr>
<td>Observations</td>
<td>460</td>
<td>368</td>
<td>207</td>
</tr>
</tbody>
</table>

Notes: Standard errors in parentheses. * indicates that the variable was added after survey start. This table presents the mean and standard deviation of the main outcomes, by endowment of minerals at the village.
<table>
<thead>
<tr>
<th>Distance</th>
<th>Mean difference: Gold - Coltan</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(t-statistic)</td>
</tr>
<tr>
<td>Distance to closest Road, 2010</td>
<td>0.642</td>
</tr>
<tr>
<td></td>
<td>(1.06)</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Distance to Natural reserve</td>
<td>-0.0254</td>
</tr>
<tr>
<td></td>
<td>(-0.01)</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Distance to Lake</td>
<td>-14.38</td>
</tr>
<tr>
<td></td>
<td>(-1.29)</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Distance to Bukavu</td>
<td>-7.357</td>
</tr>
<tr>
<td></td>
<td>(-0.87)</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Distance to Rwanda</td>
<td>-10.94</td>
</tr>
<tr>
<td></td>
<td>(-1.07)</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Distance to closest river</td>
<td>-0.343</td>
</tr>
<tr>
<td></td>
<td>(-0.32)</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Distance to closest bridge, 2010</td>
<td>-8.048***</td>
</tr>
<tr>
<td></td>
<td>(-4.16)</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Distance to closest Airport, 2010</td>
<td>-1.845</td>
</tr>
<tr>
<td></td>
<td>(-0.94)</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Accessible by car in 1999</td>
<td>0.0587</td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
</tr>
<tr>
<td></td>
<td>87</td>
</tr>
<tr>
<td>Access by moto in 1999</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
<tr>
<td></td>
<td>87</td>
</tr>
<tr>
<td>Access to phone coverage in 1999</td>
<td>-0.0190</td>
</tr>
<tr>
<td></td>
<td>(-0.38)</td>
</tr>
<tr>
<td></td>
<td>87</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the resulting t-statistics from a t-test of the difference in levels between gold and coltan villages. A positive t-statistic indicates a larger average value in gold villages.
Table 3: Effects of price shocks, conquests

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.25*</td>
<td>0.61***</td>
<td>0.26*</td>
<td>0.30**</td>
<td>0.60**</td>
<td>0.59**</td>
<td>0.60**</td>
<td>0.30**</td>
<td>0.66***</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.14)</td>
<td>(0.14)</td>
<td>(0.14)</td>
<td>(0.24)</td>
<td>(0.24)</td>
<td>(0.25)</td>
<td>(0.14)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>Coltan X Coltan Price X Far</td>
<td>-0.58***</td>
<td>-0.57*</td>
<td>-0.56*</td>
<td>-0.57*</td>
<td>-0.66**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.27)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold X Gold Price</td>
<td>-0.01***</td>
<td>0.00</td>
<td>-0.01***</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold X Gold Price X Far</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congolese Army, Village</td>
<td>-0.06**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.12***</td>
<td>-70.08</td>
<td>0.12***</td>
<td>0.15*</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.05</td>
<td>0.09***</td>
<td>0.09***</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(74.18)</td>
<td>(0.03)</td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.08)</td>
<td>(0.09)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,342</td>
<td>1,342</td>
<td>1,342</td>
<td>732</td>
<td>732</td>
<td>678</td>
<td>732</td>
<td>244</td>
<td>244</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.21</td>
<td>0.33</td>
<td>0.22</td>
<td>0.34</td>
<td>0.35</td>
<td>0.29</td>
<td>0.37</td>
<td>0.68</td>
<td>0.69</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Village FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Region*Year FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Distance to road</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results on the baseline specification for a dummy indicating presence of attempted conquests as dependent variable. Attempted conquests are defined as violent events in which an armed group engages in fighting against an other group with the aim to acquire the territorial control of the village. Standard errors are clustered at the level of the village. The variable Coltan Price is the US price of columbite-tantalum (coltan), in US$ per kilogram, and the variable Coltan is a dummy indicating whether the village is endowed with coltan, as collected in the survey. Column (1) presents the result of the baseline specification; column (2) adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample; column (3) adds a control for the Congolese army. Columns (4)-(9) restrict the observations to the years of the Second Congo War (1998-2003); column (4) presents the baseline specification; column (5) adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample; column (6) controls for the closest distance to the road interacted with year dummies; column (7) adds region*year fixed effects. Columns (8) and (9) restrict the years to 1999 and 2000 (the year preceding the coltan shock and the year of the coltan shock); column (8) presents the baseline specification, and column (9) adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample.
Table 4: Effects of price shocks, pillages

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coltan X Coltan Price</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.10</td>
<td>0.13</td>
<td>0.13</td>
<td>0.06</td>
<td>-0.67</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.22)</td>
<td>(0.15)</td>
<td>(0.16)</td>
<td>(0.24)</td>
<td>(0.24)</td>
<td>(0.24)</td>
<td>(0.93)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>Coltan X Coltan Price X Far</td>
<td>-0.00</td>
<td>-0.08</td>
<td>-0.08</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.33)</td>
<td>(0.34)</td>
<td>(0.34)</td>
<td>(0.34)</td>
<td>(0.34)</td>
<td>(0.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold X Gold Price</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>-0.01</td>
<td>-0.00</td>
<td>-0.00</td>
<td>-0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold X Gold Price X Far</td>
<td>-0.00</td>
<td>-0.03</td>
<td>-0.01</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congolese Army, Village</td>
<td>0.13***</td>
<td>0.13***</td>
<td>0.14***</td>
<td>0.17</td>
<td>-0.03</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.10</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.05)</td>
<td>(0.03)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.12)</td>
<td>(0.17)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,342</td>
<td>1,342</td>
<td>1,342</td>
<td>732</td>
<td>732</td>
<td>678</td>
<td>732</td>
<td>67</td>
<td>244</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.14</td>
<td>0.15</td>
<td>0.14</td>
<td>0.26</td>
<td>0.27</td>
<td>0.26</td>
<td>0.29</td>
<td>0.82</td>
<td>0.55</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Village FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Region*Year FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Distance to road</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results on the baseline specification for a dummy indicating presence of pillages as dependent variable. Attempted conquests are defined as violent events in which an armed group engages in fighting against another group with the aim to acquire the territorial control of the village. Standard errors are clustered at the level of the village. The variable Coltan Price is the US price of columbite-tantalum (coltan), in US$ per kilogram, and the variable Coltan is a dummy indicating whether the village is endowed with coltan, as collected in the survey. Column (1) presents the result of the baseline specification; column (2) adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample; column (3) adds a control for the Congolese army. Columns (4)-(9) restrict the observations to the years of the Second Congo War (1998-2003); column (4) presents the baseline specification; column (5) adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample; column (6) controls for the closest distance to the road interacted with year dummies; column (7) adds region*year fixed effects. Columns (8) and (9) restrict the years to 1999 and 2000 (the year preceding the coltan shock and the year of the coltan shock); column (8) presents the baseline specification, and column (9) adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample.
Table 5: Effects of price shocks, new stationary bandit

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coltan X Coltan Price</td>
<td>0.47***</td>
<td>0.81***</td>
<td>0.46***</td>
<td>0.47***</td>
<td>0.82***</td>
<td>0.83***</td>
<td>0.79***</td>
<td>0.51**</td>
<td>0.69**</td>
</tr>
<tr>
<td>(0.17)</td>
<td>(0.25)</td>
<td>(0.17)</td>
<td>(0.18)</td>
<td>(0.28)</td>
<td>(0.27)</td>
<td>(0.30)</td>
<td>(0.21)</td>
<td>(0.31)</td>
<td></td>
</tr>
<tr>
<td>Coltan X Coltan Price X Far</td>
<td>-0.67**</td>
<td>-0.65*</td>
<td>-0.49</td>
<td>-0.62</td>
<td>-0.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.33)</td>
<td>(0.35)</td>
<td>(0.35)</td>
<td>(0.37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold X Gold Price</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.05*</td>
<td>-0.08*</td>
<td>-0.08*</td>
<td>-0.09**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold X Gold Price X Far</td>
<td>-0.00</td>
<td>0.06</td>
<td>0.10</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.00)</td>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congolese Army, Village</td>
<td>0.22***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.02</td>
<td>-0.22***</td>
<td>-0.03*</td>
<td>0.17</td>
<td>0.15</td>
<td>0.11</td>
<td>0.17</td>
<td>0.16***</td>
<td>0.23</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.04)</td>
<td>(0.32)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,342</td>
<td>1,342</td>
<td>1,342</td>
<td>732</td>
<td>732</td>
<td>678</td>
<td>732</td>
<td>244</td>
<td>244</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.14</td>
<td>0.16</td>
<td>0.17</td>
<td>0.23</td>
<td>0.25</td>
<td>0.27</td>
<td>0.26</td>
<td>0.52</td>
<td>0.52</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Village FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Region*Year FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Distance to road</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results on the baseline specification for the dummy entry of a new stationary bandit as dependent variable. Attempted conquests are defined as violent events in which an armed group engages in fighting against an other group with the aim to acquire the territorial control of the village. The variable Coltan Price is the US price of columbite-tantalum (coltan), in US$ per kilogram, and the variable Coltan is a dummy indicating whether the village is endowed with coltan, as collected in the survey. Standard errors are clustered at the level of the village. Column (1) presents the result of the baseline specification; column (2) adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample; column (3) adds the control for the Congolese army. Columns (4)-(9) restrict the observations to the years of the Second Congo War (1998-2003); column (4) presents the baseline specification; column (5) adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample; column (6) controls for the closest distance to the road interacted with year dummies; column (7) adds region*year fixed effects. Columns (8) and (9) restrict the years to 1999 and 2000 (the year preceding the coltan shock and the year of the coltan shock); column (8) presents the baseline specification, and column (9) adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample.
Table 6: Effects of the coltan shock, all outcomes

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) Stationary Bandit</th>
<th>(2) Stationary Bandit</th>
<th>(3) Village Vacated</th>
<th>(4) Village Vacated</th>
<th>(5) Village Taxation</th>
<th>(6) Village Taxation</th>
<th>(7) Acquisition of Firearms</th>
<th>(8) Acquisition of Firearms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coltan X Coltan Price</td>
<td>0.24*</td>
<td>0.54***</td>
<td>0.08*</td>
<td>0.00</td>
<td>0.27*</td>
<td>0.56**</td>
<td>0.23**</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.20)</td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.15)</td>
<td>(0.22)</td>
<td>(0.10)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Coltan X Coltan Price X Far</td>
<td>-0.57**</td>
<td>0.15*</td>
<td>-0.54*</td>
<td>-0.00</td>
<td></td>
<td></td>
<td>-0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.09)</td>
<td>(0.30)</td>
<td>(0.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.53***</td>
<td>0.52**</td>
<td>-0.00</td>
<td>0.05</td>
<td>0.44***</td>
<td>0.48**</td>
<td>0.02</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.20)</td>
<td>(0.01)</td>
<td>(0.07)</td>
<td>(0.03)</td>
<td>(0.22)</td>
<td>(0.02)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Observations</td>
<td>234</td>
<td>234</td>
<td>244</td>
<td>244</td>
<td>234</td>
<td>234</td>
<td>234</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.85</td>
<td>0.86</td>
<td>0.52</td>
<td>0.53</td>
<td>0.84</td>
<td>0.85</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Village FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *** p < 0.01, ** p < 0.05, * p < 0.1. Years are restricted to 1999 and 2000 (the year preceding the coltan shock and the year of the coltan shock). This table presents the results on the baseline specification for the following outcomes: presence of a stationary bandit, whether the village transitions from having a stationary bandit to not having one, whether an armed group is collecting taxes in the village, and a dummy variable indicating whether an armed group acquires weapons in this village. A stationary bandit is defined as an armed group that is settled in the village for at least 6 months and has the monopoly of violence in the village. The variable Coltan Price is the US price of columbite-tantalum (coltan), in US$ per kilogram, and the variable Coltan is a dummy indicating whether the village is endowed with coltan, as collected in the survey. Standard errors are clustered at the level of the village. Columns (1), (3), (5), (7), report the results considering the baseline specification, while columns (2), (4), (6), (8), adds the interaction of the main regressor (Coltan X Coltan Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample.
### Table 7: Effects of the gold shock, all outcomes

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stationary Bandit</td>
<td>Stationary Bandit</td>
<td>Village Vacated</td>
<td>Village Vacated</td>
<td>Village Taxation</td>
<td>Village Taxation</td>
<td>Acquisition of Firearms</td>
<td>Acquisition of Firearms</td>
</tr>
<tr>
<td>Gold X Gold Price</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00*</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Gold X Gold Price X Far</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.70***</td>
<td>0.91***</td>
<td>-0.01</td>
<td>-0.15***</td>
<td>0.67***</td>
<td>0.84***</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.07)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.07)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Observations</td>
<td>936</td>
<td>936</td>
<td>976</td>
<td>976</td>
<td>936</td>
<td>936</td>
<td>936</td>
<td>936</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.60</td>
<td>0.61</td>
<td>0.11</td>
<td>0.13</td>
<td>0.58</td>
<td>0.59</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Village FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. Years are restricted to 2001-2008 (the years following the coltan shock until the Kimia II intervention). This table presents the results on the baseline specification for the following outcomes: presence of a stationary bandit, whether the village transitions from having a stationary bandit to not having one, whether an armed group is collecting taxes in the village, and a dummy variable indicating whether an armed group acquires weapons in this village. A stationary bandit is defined as an armed group that is settled in the village for at least 6 months and has the monopoly of violence in the village. The variable Gold Price is the US price of gold, in US$ per kilogram, and the variable Gold is a dummy indicating whether the village is endowed with gold, as collected in the survey. Standard errors are clustered at the level of the village. Columns (1), (3), (5), (7), report the results considering the baseline specification, while columns (2), (4), (6), (8), adds the interaction of the main regressor (Gold X Gold Price) with a dummy indicating whether the village distance to the closest airport is in above the median of the sample.
Table 8: Armed groups’ taxes

<table>
<thead>
<tr>
<th></th>
<th>Gold</th>
<th>Coltan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mining: per mineral output</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>(0)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Mining: per labor input</td>
<td>0.18</td>
<td>0.13</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Poll taxes</td>
<td>0.54</td>
<td>0.52</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Market transactions</td>
<td>0.19</td>
<td>0.13</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Transit of persons</td>
<td>0.19</td>
<td>0.25</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Village mill</td>
<td>0.12</td>
<td>0.07</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>

Notes: This table presents the taxation results. The numbers indicate the proportion of villages in which any of the following taxes was collected by an armed group: tax on mineral output (first line), tax on labor inputs at the mine (second line), tax on households (third line), tax on sale of cassava at the local market (fourth line), tax on transit of persons (fifth line), tax on village mill (sixth line).

Table 9: Price shocks and taxes

<table>
<thead>
<tr>
<th>VARIABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold X Gold Price</td>
</tr>
<tr>
<td>Coltan X Coltan Price</td>
</tr>
<tr>
<td>Coltan</td>
</tr>
<tr>
<td>Constant</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Output tax roadblocks</th>
<th>Entry tax</th>
<th>Food sales tax</th>
<th>Poll tax</th>
<th>Transit tax</th>
<th>Mill tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td></td>
</tr>
<tr>
<td>Gold X Gold Price</td>
<td>0.00122</td>
<td>0.00142</td>
<td>0.00644*</td>
<td>-0.00617***</td>
<td>0.00222</td>
<td></td>
</tr>
<tr>
<td>(0.00336)</td>
<td>(0.00252)</td>
<td>(0.00330)</td>
<td>(0.00299)</td>
<td>(0.00162)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coltan X Coltan Price</td>
<td>0.515***</td>
<td>0.393***</td>
<td>-0.00768</td>
<td>0.337**</td>
<td>0.0802</td>
<td>0.0575</td>
</tr>
<tr>
<td>(0.105)</td>
<td>(0.0927)</td>
<td>(0.0949)</td>
<td>(0.138)</td>
<td>(0.0925)</td>
<td>(0.0542)</td>
<td></td>
</tr>
<tr>
<td>Coltan</td>
<td>0.110**</td>
<td>-0.0834</td>
<td>-0.0377</td>
<td>0.0681</td>
<td>-0.0700</td>
<td>-0.00985</td>
</tr>
<tr>
<td>(0.0427)</td>
<td>(0.0691)</td>
<td>(0.0575)</td>
<td>(0.0757)</td>
<td>(0.0689)</td>
<td>(0.0348)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0259*</td>
<td>0.106**</td>
<td>0.118***</td>
<td>0.237***</td>
<td>0.189***</td>
<td>0.0510***</td>
</tr>
<tr>
<td>(0.0151)</td>
<td>(0.0523)</td>
<td>(0.0389)</td>
<td>(0.0502)</td>
<td>(0.0473)</td>
<td>(0.0225)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results on the baseline specification for dummies indicating whether armed groups collect a particular tax in the village as dependent variables. Standard errors are clustered at the village level. The dependent variable in column (1) is a dummy indicating whether an armed group collects a tax on output at roadblocks outside the mine. The dependent variable in column (2) is a dummy indicating whether an armed group collects a tax on labor at the mine (entry tax). The dependent variable in column (3) is a dummy indicating whether an armed group collects a tax on sale of cassava at the local market. The dependent variable in column (4) is a dummy indicating whether an armed group collects a tax on households (poll tax). The dependent variable in column (5) is a dummy indicating whether an armed group collects a tax on transit of persons in and out of the village. Finally, the dependent variable in column (6) is a dummy indicating whether an armed group collects a tax on a local mill in the village.
Table 10: Price shocks and tax levels

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output tax</td>
<td>3.08***</td>
<td>2.29</td>
<td>0.01</td>
<td>7.50***</td>
<td>8.60**</td>
<td>5.03</td>
</tr>
<tr>
<td>roadblocks</td>
<td>(0.75)</td>
<td>(2.56)</td>
<td>(0.20)</td>
<td>(2.06)</td>
<td>(3.81)</td>
<td>(5.99)</td>
</tr>
<tr>
<td>Entry tax</td>
<td>-0.14**</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.06</td>
<td>0.10</td>
<td>0.06</td>
</tr>
<tr>
<td>Food sales tax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poll tax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transit tax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mill tax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold X Gold Price</td>
<td>-1.64***</td>
<td>-2.14**</td>
<td>-0.37***</td>
<td>-0.29</td>
<td>-7.22***</td>
<td>-14.15***</td>
</tr>
<tr>
<td></td>
<td>(0.45)</td>
<td>(0.92)</td>
<td>(0.08)</td>
<td>(0.60)</td>
<td>(1.64)</td>
<td>(0.53)</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results on the baseline tobit specification for left-truncated variables indicating the level of different taxes dependent variables. Standard errors are clustered at the village level. The dependent variable in column (1) is the tax on output at roadblocks outside the mine, in US$ per kilogram. The dependent variable in column (2) is the tax on labor at the mine (US$ per entry to the mine). The dependent variable in column (3) is the tax on cassava sales at the local market, in US$ per day of sale. The dependent variable in column (4) is the tax on households (US$ per week). The dependent variable in column (5) is the tax on transit of persons in and out of the village, in US$ per passage. Finally, the dependent variable in column (6) a tax on a local mill in the village, in US$ per day of activity.

Table 11: A test of the fiscal externality channel

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>model</td>
<td>1.937e+06**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(859,293)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-609,841***</td>
<td>482,854***</td>
</tr>
<tr>
<td></td>
<td>(124,753)</td>
<td>(92,741)</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results of a tobit regression of the index of taxes on one day of labor in the mining sector on the index of daily taxes on cassava sales at the local market. The indexes are the estimated day equivalent tax in dollars in each sector. For the mining sector, I use taxes at the mine of the village, while for the agricultural sector I use taxes at the market on cassava sales. I then computed a daily sector i tax index for each sector as follows: \( I_i = \left( \hat{Y}_i \tau_i + t_i \right) p_i \), where \( \hat{Y}_i \) is the estimated average volume product per worker for one day of work, \( \tau_i \) is sector i tax on output, \( t_i \) is sector i tax levied on one day of work, and \( p_i \) is the local output price in sector i. Since I do not observe taxes on output for gold and since the tax on cassava is charged on one day of sales at the market, \( \tau_i = 0 \) in gold and agriculture. I therefore estimate production for heavy minerals, coltan and cassiterite. Based on the average production reported in all year-village observations, I imputed 20 kilograms both for cassiterite and coltan. Standard errors are clustered at the village level.
Table 12: Stationary to roving: OLS

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FDLR taxbase</td>
<td>Village taxes</td>
<td>FDLR Attack</td>
<td>Pillage</td>
<td>Pillage</td>
<td>Pillage</td>
<td>Pillage</td>
<td>Pillage</td>
</tr>
<tr>
<td>KIMIA2</td>
<td>-0.741***</td>
<td>-0.389***</td>
<td>0.278***</td>
<td>0.259***</td>
<td>0.250***</td>
<td>0.228**</td>
<td>0.196**</td>
<td>0.259***</td>
</tr>
<tr>
<td></td>
<td>(0.0936)</td>
<td>(0.116)</td>
<td>(0.0827)</td>
<td>(0.0792)</td>
<td>(0.0759)</td>
<td>(0.0893)</td>
<td>(0.0952)</td>
<td>(0.0632)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.140***</td>
<td>0.494***</td>
<td>0.0763***</td>
<td>0.0903***</td>
<td>0.0916***</td>
<td>0.0992***</td>
<td>0.120***</td>
<td>0.0903***</td>
</tr>
<tr>
<td></td>
<td>(0.00643)</td>
<td>(0.00959)</td>
<td>(0.00568)</td>
<td>(0.00544)</td>
<td>(0.00250)</td>
<td>(0.00259)</td>
<td>(0.0311)</td>
<td>(0.0125)</td>
</tr>
<tr>
<td>Observations</td>
<td>786</td>
<td>654</td>
<td>786</td>
<td>786</td>
<td>786</td>
<td>726</td>
<td>600</td>
<td>262</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.788</td>
<td>0.703</td>
<td>0.335</td>
<td>0.279</td>
<td>0.320</td>
<td>0.292</td>
<td>0.290</td>
<td>0.601</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Village FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Region*Year FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Distance to road</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Accessibility</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results on the OLS regression of FDLR behavior. The regressor Post*FDLR interacts Post and FDLR. Post is a dummy variable indicating whether the observation is after 2008, and FDLR is a dummy variable indicating whether the village was controlled by the FDLR in 2008, and hence a target of the Kimia II operation. All regressions include year and village fixed effects. Standard errors are clustered at the village level. The regressors Post and FDLR are thus absorbed by the fixed effects. Columns (1)-(3) use the following variables as dependent variables, respectively: FDLR tax base (a dummy indicating whether the FDLR collects taxes in the village in that given year); Village taxes (a dummy indicating whether any group collects taxes in the village in that given year); FDLR attack (a dummy indicating whether the village received an attack by the FDLR in that given year). Columns (4)-(8) present variations to the baseline specification, using as dependent variable a dummy indicating whether the village was subject to pillages in that given year. Pillages are defined as a violent event in which an armed group enters the village in order to acquire household assets and leaves the village after. Pillages tend to be very violent and are straightforward to identify. Column (4) presents the results on pillage for the main specification; column (5) adds region*year fixed effects, column (6) adds distance to the road, interacted with year dummies; column (7) adds a time-varying control dummy indicating whether the village has access to the road; column (8) collapses all observations by treatment status and by post, thus ignoring the time-series information.
Table 13: IV estimates of RCD protection (baseline specification)

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) Poll tax</th>
<th>(2) Pillaging</th>
<th>(3) Marriages</th>
<th>(4) Activity</th>
<th>(5) Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCD</td>
<td>0.512***</td>
<td>0.0488</td>
<td>-1.205</td>
<td>0.250**</td>
<td>0.220**</td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
<td>(0.108)</td>
<td>(1.511)</td>
<td>(0.114)</td>
<td>(0.0877)</td>
</tr>
<tr>
<td>Observations</td>
<td>420</td>
<td>436</td>
<td>420</td>
<td>316</td>
<td>436</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.123</td>
<td>-0.001</td>
<td>-0.001</td>
<td>0.007</td>
<td>0.024</td>
</tr>
<tr>
<td>Number of groupIDV_mine</td>
<td>105</td>
<td>109</td>
<td>105</td>
<td>79</td>
<td>109</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Village FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results on the 2SLS regression of the main welfare outcomes on presence of the Rassemblement Congolais pour la Démocratie. The endogenous regressor RCD is a dummy variable indicating whether the RCD controlled the village in a given year. The first stage regresses RCD on an interaction of a variable indicating whether the observation is posterior to the year 2003, and a dummy indicating whether the observation is a village that was under the control of the RCD in the year 2002, the year prior to the Sun-City agreement. All regressions include year fixed effects and village fixed effects. Standard errors are clustered at the level of the village. The following outcomes are the dependent variables across columns. Poll tax is a dummy indicating whether an armed group is collecting taxes on households in a given year (column (1)). Pillaging is a dummy indicating whether an armed group pillaged the village in a given year (column (2)). Marriages indicates the number of marriages celebrated in the village in a given year (column (3)). Activity is a dummy indicating whether mining activity was recorded in the main mine of the village in a given year (column (4)). Output is a dummy indicating whether mining output was recorded in the main mine of the village in a given year (column (5)).
<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCD</td>
<td>0.220**</td>
<td>0.174*</td>
<td>0.202**</td>
<td>0.150*</td>
<td>0.0808</td>
<td>0.0995</td>
</tr>
<tr>
<td></td>
<td>(0.0877)</td>
<td>(0.0899)</td>
<td>(0.0972)</td>
<td>(0.0893)</td>
<td>(0.116)</td>
<td>(0.151)</td>
</tr>
<tr>
<td>Observations</td>
<td>436</td>
<td>436</td>
<td>436</td>
<td>436</td>
<td>400</td>
<td>348</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.024</td>
<td>0.068</td>
<td>0.067</td>
<td>0.168</td>
<td>0.126</td>
<td>0.082</td>
</tr>
<tr>
<td>Number of groupIDV_mine</td>
<td>109</td>
<td>109</td>
<td>109</td>
<td>109</td>
<td>100</td>
<td>87</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Village FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Region*Year FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Distance to road</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Accessibility</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Mineral controls</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results on the 2SLS regression of the main welfare outcomes on presence of the Rassemblement Congolais pour la Démocratie. The endogenous regressor RCD is a dummy variable indicating whether the RCD controlled the village in a given year. The first stage regresses RCD on an interaction of a variable indicating whether the observation is posterior to the year 2003, and a dummy indicating whether the observation is a village that was under the control of the RCD in the year 2002, the year prior to the Sun-City agreement. All regressions include year fixed effects and village fixed effects. Standard errors are clustered at the level of the village. Column (1) presents the baseline 2SLS regression. Column (2) controls for the interaction between a dummy indicating whether the village is endowed with coltan and the US price of coltan in US$ per kilogram (Coltan X Coltan Price). Column (3) adds the interaction between a dummy indicating whether the village is endowed with gold and the US price of coltan in US$ per kilogram (Gold X Gold Price). Column (4) includes region*year fixed effects to the baseline specification. Column (5) controls for the distance to the road, interacted with year dummies.
Figure 1: Timing of operations: pillages vs. conquest attempts

Notes: This figure draws on the attacks level dataset collected in this survey. In total, there are 600 attacks. The figure plots the distribution of the recorded hour of the attack. The horizontal axes indicate the hour of the attack in 24h format. The vertical axes indicate the mass of observations. The left quadrant plots the distribution for recorded pillages, and the right quadrant plots the distribution for recorded conquest attempts. The red, vertical lines indicate the hours at which the sun rises (left line) and sets (right line). In the outer of these lines the village is in the dark.

Figure 2: Output realization and amount hidden from taxes

Notes: This figure represents the relationship between the volume of output hidden in sector $i$ ($H_i$), and the realized output in sector $i$ ($\tilde{H}_i$) derived from the household’s first order conditions. If the realized labor volume productivity is below a certain threshold, the household is at a corner solution and hides all output. If the realized labor volume productivity is above the threshold, the volume hidden is an interior solution. Any additions to the realized volume productivity do not increase the volume hidden by the household.
Notes: This map presents the spatial location of the survey villages across Sud Kivu. Thick lines indicate frontiers between Provinces, dashed lines indicate frontiers between Territoires. Dark green indicates a natural reserve (forest). Blue indicates the lake Kivu. The villages endowed with coltan only are indicated in black circles. The villages endowed with coltan and gold are indicated as black dots with golden circles. Villages endowed with cassiterite only are indicated with orange dots, and agricultural villages are indicated as crosses. In each Territoire (Shabunda, Mwenga, Walungu, Kabare, Kalehe) the survey visited the population of coltan villages that were accessible to the survey teams. In addition, it visited a sample of gold villages randomly selected within the Territoire. In addition to this, in the Territoire of Kalehe, the survey visited all mining villages, as well as a matched sample of agricultural villages. To match agricultural villages to the population of mining villages in Kalehe prior to the survey, I matched multiple data sources, including SAESSCAM (Congolese institute of mining), the Institute of Peace (Belgium), and data I collected with my own qualitative fieldwork in May-July 2012 with mining traders. After compiling the population of mining villages of Kalehe, I then used their geographic coordinates to assign to each a distance to: the lake, the natural reserve (forest), Bukavu, Goma, Rwanda, and the road. Since mining villages are likely to be more remote by the fact that they are more valuable and populations self-select into locations, I matched the resulting population of mining villages on the population of villages based on the Mahalanobis metric on these variables. To obtain geo-locations of the population of villages prior to the survey, I used publicly available data from Référentiel Géographique Commun (2010), that I compiled with a dataset collected by the International Rescue Committee and Humphreys, Sanchez de la Sierra, and Van der Windt (2012, 2013a), who mapped all villages in most Territoires of Sud Kivu and collected geo-locations for each.
Figure 4: Control by armed groups

Notes: This figure graphs the number of villages in the sample under the control of a given armed group. The vertical axis represents the number of villages, and the horizontal axis represents the years. Labels for armed groups are included in the legend. The data fits well-known historical phases of the conflict and post-conflict periods. The dashed vertical lines indicate the start and end of the occupation by the Rassemblement Congolais pour la Démocratie.

Figure 5: Activities of stationary bandits

Notes: This figure graphs the share of villages in the sample in which an armed group coded as stationary bandit implements the following activities: is a stationary bandit, collects taxes, administers justice, administers the village. Armed groups were coded as stationary bandits when they had the monopoly of violence at the village for at least 6 months. The vertical axis represents the share of villages, and the horizontal axis represents the years. Labels for activities are included in the legend.
Figure 6: Distribution of protected and unprotected episodes in the sample

Notes: The left figure plots the distribution of the duration for all episodes in the sample in which a village was under the control of a stationary bandit. The right figure plots the distribution of the duration for all episodes in the sample in which a village was under the control of no stationary bandit. For both figures, the vertical axes represent the density and the horizontal axes indicate the number of years.

Figure 7: World prices of coltan and gold, by year

Notes: This figure plots the yearly average price of gold and coltan, in US$ per kilogram as recorded in the US market. The left vertical axis indicates the price of coltan and the right vertical axis indicates the price of gold. The horizontal axis indicates the years. Source: United States Geological Survey (2010).
Figure 8: Effect of the coltan shock on attacks on mining camps

Notes: This figure plots the proportion of villages, whose main mine was attacked, by year. The left quadrant plots this relationship for gold villages, and the right quadrant for coltan villages. The thick line indicates the proportion of villages in which an attack took place. The right vertical axis indicates the local price of gold in US$ per gram. The dashed line represents the price of gold and the vertical line indicates the timing of the coltan shock.

Figure 9: Effect of the coltan shock on Conquest attempts on villages

Notes: This figure plots the proportion of villages that experienced a conquest attempt, by year. The left quadrant plots this relationship for gold villages, and the right quadrant for coltan villages. The thick line indicates the proportion of villages in which an attack took place. The right vertical axis indicates the local price of gold in US$ per gram. The dashed line represents the price of gold and the vertical line indicates the timing of the coltan shock.
Figure 10: First main effect (P1): summary of effects of the coltan shock

Notes: This graph summarizes the results on the coltan price shock (Prediction P1). At the time of the coltan shock, villages were either occupied by an armed group (stationary bandit present) or not (Stationary bandit absent). For an occupied village, armed groups may attempt a conquest ("Conquest attempt") or not ("No conquest attempt"). The attempt may be successful, in which case the armed group emerges as the new stationary bandit ("Entry") or it may fail ("No Entry"). If there is no conquest attempt, the stationary bandit present at the village may also decide to leave ("Exit") or remain as stationary bandit ("Stay"). In the lower part of the diagram, if no stationary bandit is present, an armed group may emerge as the village stationary bandit ("Entry") or the village may remain unoccupied ("No Entry").
Figure 11: Second main effect (P2): summary of effects of the coltan shock

Notes: This graph summarizes the results on the gold price shock (Prediction P2). At the time of the gold shock, villages were either occupied by an armed group (stationary bandit present) or not (Stationary bandit absent). For an occupied village, armed groups may attempt a conquest ("Conquest attempt") or not ("No conquest attempt"). The attempt may be successful, in which case the armed group emerges as the new stationary bandit ("Entry") or it may fail ("No Entry"). If there is no conquest attempt, the stationary bandit present at the village may also decide to leave ("Exit") or remain as stationary bandit ("Stay"). In the lower part of the diagram, if no stationary bandit is present, an armed group may emerge as the village stationary bandit ("Entry") or the village may remain unoccupied ("No Entry").
Figure 12: Main outcomes, matching results

Notes: These figures present the results from the tests matching coltan villages on non coltan villages based on the Mahalanobis metric. As matching variables, I use distance to airports and distance to roads, as well as upper level administrative divisions (Territoire). Since airport areas might be more urbanized, and urbanization might affect in unknown ways the behavior of armed groups, the procedure also considers distance to roads. Finally, to control for shocks that commonly affect regions (Territoires), I also match on Territoire. This procedure is thus equivalent to conditional differences-in-differences within calipers defined by administrative divisions (Cochran and Rubin, 1973). Using this method, I obtain matched pairs for each year. I then regress the difference in outcomes between the coltan village of the pair and the non coltan village of the pair on year dummies. The graphs present the coefficient on the year dummies, and their 90% confidence intervals. The black line shows the yearly coltan coefficients from an OLS regression of the marched pairs on year dummies. Dotted lines represent the confidence intervals. Standard errors are clustered at the pair level.
Notes: These figures present the results from the randomization inference exercise. The upper figure plots the distribution of coefficients on the main interaction Coltan X Coltan Price obtained by randomly assigning prices of coltan to different years. The vertical line indicates the value of the coefficient estimated using real prices. The resulting p-value is 0.065. The bottom figure plots the distribution of coefficients on the variable Coltan X Coltan Price X d, where d is a dummy indicating whether the village distance to the closest airport is above the median distance to the closest airport in the sample villages. The resulting p-value is 0.09.
Figure 14: Armed group output tax levels, by year and mineral endowment

Notes: This figure plots the average output tax level on the mineral, collected at the roadblock at mine exit, per year. The left quadrant plots this relationship for gold villages, and the right quadrant plots this relationship for coltan villages. The solid line indicates the specific tax rate on output. The tax rate is converted to US$ per kilogram and its axis is reported on the left vertical axis. The local price of coltan, represented as a dashed line, is expressed in US$ per kilogram, and reported on the right vertical axis. Output taxes are collected based on the fee that miners have to pay for each bag of coltan.

Figure 15: Armed group taxes on mining labor, by year and mineral endowment

Notes: This figure plots the average level for the tax on labor (fee to enter the mine), collected at the entry at mine, per year. The left quadrant plots this relationship for gold villages, and the right quadrant plots this relationship for coltan villages. The solid line indicates the tax level on mining labor, captured at the entry of the mine. It is the fee paid for one day of work. The tax on labor is in US$ per kilogram and the values are reported on the left vertical axis. The local price of coltan, represented as a dashed line, is expressed in US$ per kilogram, and its values are reported on the right vertical axis.
Notes: The first two figures show the proportion of villages controlled by the different armed groups between 2002 and 2012. The sample was split into villages targeted by the Kimia II operations (left) and villages that were not targeted (right). I define a targeted village by whether it was under the control of the FDLR in 2008, since Kimia II targeted all FDLR villages. The left quadrant presents the proportion of villages under the control of each of the indicated armed groups, for villages that were not controlled by the FDLR in 2008, prior to Kimia II. The right quadrant shows the proportion of villages under the control of each of the indicated armed groups, for villages that were controlled by the FDLR in 2008, prior to Kimia II. The line marked with squares indicates the proportion of villages controlled by the FDLR, the line marked with diamonds indicates the proportion of villages under the control of the Congolese army. The bottom figures are divided with the same logic between control and treated villages. However, the outcomes are the proportion of villages in which a pillage by the FDLR took place. The left axis indicates this proportion.
Appendices

A Mathematical Appendix

A.1 Households first order conditions

There are $N$ first order conditions for $e_i$, $N$ first order conditions for $e_i^H$, $N$ first order conditions for $H_i$ conditioned on $\tilde{\alpha}_i e_i$ plus the $N + 1$ first order conditions from the constraints. The Kuhn-Tucker conditions with respect to $H_i$, $i = 1, \ldots, N$ give:

$$\forall \tilde{\alpha}_i, \ u'(.)[\tau_i p_i - G^i(H_i, p_i)]\mu_i = 0, \ i = 1, \ldots, N$$

and

$$\tau_i p_i - G_1(H_i, p_i) \geq 0 \ i = 1, \ldots, N$$

An interior solution is therefore given by $H_i < \tilde{\alpha}_i e_i$ and $H_i > 0$, hence $\mu_i = 0$ and:

$$\tau_i p_i = G_1(H_i, p_i)$$
Figure 18: Heterogeneous effects of price shocks, by expectations

Notes: This figure plots the transitions of entry and exit by armed groups in the survey villages, in and out of stationary bandit. For all figures the horizontal axis indicates the year and the vertical axis on the right indicates the proportion of villages in which an event takes place. The vertical axis in the middle indicates the price of coltan per gram, in US$. The left vertical axis indicates the gold price per kilogram, in US$. Inside each figure, the dashed line represents the local price of coltan, the dotted yellow line indicates the world price of gold. The black solid line indicates the outcome of interest. Turning to the two figures above, the outcome of interest is the emergence of new stationary bandit in the village (“Entry”). The left quadrant plots the emergence of new stationary bandits on the year for villages in which respondents today reported that they expected the price rise to be temporary. The corresponding plot for villages that reported the price rise to be permanent is presented on the right. These figures suggest that armed groups entered in response to the coltan shock mostly in villages in which populations thought it would be a permanent shock. Turning to the lower figures, the outcome of interest is the departure of any stationary bandit from the village. The left quadrant plots this variable on years for villages in which the price drop was expected to be permanent and the right quadrant plots the variable for villages in which the price drop was expected to be temporary. In 2002, armed groups departed mostly from villages where there is consensus that populations thought the price drop would be permanent.
Similarly, an interior solution for \( e^H \) is given by \( e^H < e_i \) and \( e^H > 0 \):

\[
t_i = E_1(e^H, p_i)
\]

The FOC with respect to \( e_i \), \( i = 1, \ldots, N \) give:

\[
\int (u'(.)[(1 - \tau_i)p_i\hat{a}_i - t_i] + \mu_i)dF(\alpha) - \lambda - c'(e_i) = 0, \ i = 1, \ldots, N
\]

for sectors where the hiding constraint is binding, and:

\[
\int (u'(.)[(1 - \tau_i)p_i\hat{a}_i - t_i])dF(\alpha) - \lambda - c'(e_i) = 0, \ i = 1, \ldots, N
\]

for sectors where it is not.

### A.2 The impossibility to tax gold output

In this discussion, there is only one sector. I drop the sector identifiers. Let \( G(H, p) = k H^2 \). The parameter \( k \) captures the cost of hiding output. Let \( c(e) = \frac{e^2}{2} \) and the utility function be linear in consumption. Labor supply in this form is isoelastic, and the optimal tax is:

\[
\tau^* = \frac{1}{2} \frac{\alpha^2}{\alpha^2 + \frac{1}{k}}
\]

The optimal tax is larger the higher is the cost to conceal an additional unit of output. At the limit, if output cannot be concealed, and \( k = +\infty \), \( \tau^* = \frac{1}{2} \). This is because labor supply is isoelastic and always equal to 1, which can be seen from the inverse elasticities rule. If \( k \in R^+ \), \( 0 < \tau^* < \frac{1}{2} \). When \( \alpha \) is large \( \tau^* \) approaches its upper bound, and when \( \alpha \) is small, \( \tau^* \) tends to 0. Since \( \alpha_g \approx 0 \), the optimal tax in a village where the only sector is gold will be low. Furthermore, if \( \alpha_g \) is sufficiently low so that \( \alpha_g L < H \) (assumption G1), the choice of tax is irrelevant and always raises no revenues, since output is always hidden. Turning to assumption G2, the tax revenue can be written as:

\[
R = \frac{1}{2} \frac{\alpha^2 p^2}{\alpha^2 + \frac{1}{k}} \left( \alpha^2 (1 - \tau) - \frac{\tau}{k} \right)
\]

\[
= \frac{1}{4} \frac{\alpha^4 p^2}{\alpha^2 + \frac{1}{k}}
\]

which is strictly increasing in \( \alpha \). Therefore, \( \exists \alpha \) s.t. \( \forall \alpha < \alpha \),

\[
R(\alpha) \leq F_g, \ \tau^* = 0
\]
and \( \forall \alpha < \alpha \),

\[ R(\alpha) > F_g, \quad \tau^* > 0 \]

where \( F_g \) was defined as the fixed cost of levying an output tax in the gold sector.

### A.3 Proof of proposition 2

The household’s and the bandit’s programs are now, respectively:

\[
\max_{e, e_{\text{H}}, H} (1 - \tau) p e + \tau p H - t (e - e_{\text{H}}) - \frac{1}{2} e^2 - \frac{s}{2} e_{\text{H}}^2 - \frac{p}{2} H^2
\]

\[
\max_{t, \tau} \tau p (\alpha e - H) + t (e - e_{\text{H}})
\]

The optimal taxes are \( \tau^* = \frac{1}{2} \frac{1}{1 + s} \) and \( t^* = \frac{\alpha p}{2} \frac{c}{1 + c} \frac{s}{1 + s} \), where \( s = \frac{1 + c}{\alpha^2 h} \). Applying the Envelope theorem to the bandit’s objective function, it follows that:

\[
\frac{\partial V}{\partial p} = \frac{\alpha^2 p}{(1 + s)^2} \left( \frac{1}{2} + s + \frac{s^2}{2} \frac{c}{1 + c} \right)
\]

For gold, there is no tax on output, we can assume \( s = 0 \), and therefore:

\[
\frac{\partial V}{\partial p_g} = \frac{\alpha^2 p}{2} \frac{c}{1 + c}
\]

We then have:

\[
\frac{\partial V}{\partial p_c} - \frac{\partial V}{\partial p_g} = \frac{\alpha^2 p_c}{(1 + s)^2} \left( \frac{1}{2} + s + \frac{s^2}{2} \frac{c}{1 + c} \right) - \frac{\alpha p}{2} \frac{c}{1 + c}
\]

\[
= \frac{1}{2} \frac{\alpha^2 p_c}{(1 + s)^2} (1 + 2 s) \left( 1 - \frac{c}{1 + c} \frac{\alpha^2 p}{\alpha^2 p_c} \right)
\]

This is positive whenever \( \alpha_g = \alpha_c \) and \( p_g = p_c \). It follows that:

\[
\frac{\partial V}{\alpha^2 p_c} - \frac{\partial V}{\alpha^2 p_g} > \frac{1}{2} \frac{1 + 2 s}{(1 + s)^2} \frac{1 + c}{1 + c}
\]

> 0

### A.4 Optimal tax properties

This section derives some of the results introduced in Section 3.4.

**Price shocks and the optimal tax**

If labor supply is isoelastic, a price shock has no effect on the optimal tax. To see when price
changes increase the optimal tax, suppose there is only one sector, that there are no labor hiding or output concealing activities (for instance, when the cost of both activities is infinite), that the utility function is separable in consumption, \((1 - \tau) p e^\alpha\), and labor disutility, \(c(e)\). From the household’s problem, the optimal labor supply is given by \(c'(e^*) = (1 - \tau)p\alpha\). Let \(g(e) = c'(e)\), the optimal tax formula can be rewritten as:

\[
\tau^* = \frac{g^{-1}((1 - \tau)p\alpha)}{g^{-1}((1 - \tau)p\alpha)p\alpha}
\]

An increase in the output price \(p\) increases the optimal tax if and only if:

\[
\frac{g^{-1'}}{g^{-1}}(1 - \tau)p\alpha > 1 + \frac{g^{-1''}}{g^{-1'}}p
\]

Hence a price increases the optimal tax when \(g^{-1'}\) is sufficiently large relative to \(g^{-1}\). This will be the case when \(c''(e)\) is sufficiently small. In this case, the elasticity of labor supply decreases with the level of the price and the optimal tax increases.

**Poll tax**

The poll tax is a tax on households, where the activity being taxed is living permanently in the village. Since migration is costly, the elasticity of living permanently in the village is low, and the optimal tax on living in the village may be positive. This can be seen directly from the optimal tax formula.

**Incentives to increase the cost of concealing output**

Suppose there is only one sector, I hence drop the sector identifiers. Let \(G(H, p) = kH^2\). The parameter \(k\) captures the sector cost of hiding output. Let \(c(e) = \frac{e^2}{2}\) and the utility function be linear in consumption. Labor supply in this form is isoelastic, and the optimal tax is:

\[
\tau^* = \frac{1}{2} \frac{\alpha^2}{\alpha^2 + \frac{1}{k}}
\]

The optimal tax increases in \(k\). Hence the costlier it is to conceal output, the higher the optimal tax. Furthermore, the stationary bandit’s revenue at the optimum is:

\[
R = \frac{1}{2} \frac{\alpha^2 p^2}{\alpha^2 + \frac{1}{k}} \left(\alpha^2 (1 - \tau) - \frac{\tau}{k}\right)
= \frac{1}{4} \frac{\alpha^2 p^2}{\alpha^2 + \frac{1}{k}}
\]

Hence, \(R\) is increasing in \(k\). If the bandit can manipulate \(k\) by using power to increase the cost
of concealing output, he will do so since this raises his tax revenue.

**Mix of input and output taxes**

Proposition: If either both the bandit and households are risk averse, or if both output and input can be hidden, the optimal contract at extraction sites will be characterized by a mix of input and output taxes. Furthermore, in the absence of risk aversion, but with imperfect monitoring of inputs and outputs, the ratio of output to input taxes will be characterized by the following relationship:

$$\tau_i^* = t_i^* \left( \alpha_i e_i - H_i \right) \frac{G_{i1}}{p_i E_{i1}}$$

Under either risk aversion or imperfect observability (or both), if stationary bandits are offering optimal contracts, they should provide a mix of input and output taxes. First, because the optimal contract is an insurance contract. Indeed, by levying a tax on output, the bandit is absorbing the risk stemming from output realizations. On the contrary, there is no risk transfer associated with a task on labor inputs, since labor supply is certain. The optimal contract will thus be characterized by a contract in which risk is bared disproportionately by the actor with lower risk aversion. Second, even in the absence of risk aversion, when either outputs or inputs are imperfectly monitored, the optimal contract will be characterized by a mix of input and output taxes. Sheltering activities generate losses through two channels: since they consist of hiding, they decrease the surplus available to tax; in addition, since they are costly, they decrease the overall surplus.

Proof that the optimal tax on output is positive when households are risk averse. Consider the one sector problem without sheltering activities:

$$\max_{t, \tau} (t + p \bar{\alpha} \tau) e = R$$

s.t. \( e = \arg\max \{ Eu ((1 - \tau) p \bar{\alpha} e - te) - c(e) \} \)

where \( \bar{\alpha} \) is the mean of the volume productivity distribution. I use the fact that the constraint is a first order condition on the producer’s objective function. From the Lagrangian first order conditions:

$$\bar{\alpha} = \frac{E \left[ u'' ((1 - \tau) p \bar{\alpha} e - te) ((1 - \tau) p \bar{\alpha} - t) \hat{\alpha} e + \hat{\alpha} u' ((1 - \tau) p \bar{\alpha} e - te) \right]}{E \left[ u'' ((1 - \tau) p \bar{\alpha} e - te) ((1 - \tau) p \bar{\alpha} - t) e + u' ((1 - \tau) p \bar{\alpha} e - te) \right]}$$

Exploiting the definition of the Coefficient of relative risk aversion, this expression is equivalent to:

$$E \left[ u' ((1 - \tau) p \bar{\alpha} e - t) (\hat{\alpha} - \bar{\alpha}) \right] = 0$$
Which implies:

\[ \text{cov} \left( u' \left( (1 - \tau) \tilde{p}\tilde{\alpha}e - t \right) (\tilde{\alpha}) \right) = 0 \]

This obtains only for \( \tau = 1 \). Obviously, the optimal contract is characterized by the bandit (principal) absorbing all the risk from production. Inserting \( \tau = 1 \) on the program leads to

\[ u'' (-te) = \frac{-1}{\lambda} \]

where \( \lambda \) is the lagrange multiplier on the producer maximization incentive compatibility constraint. This implies \( t < 0 \). The optimal contract if the bandit was risk neutral is one in which the bandit provides payment to producers in order to work that does not depend on output (a wage), and keeps all output. This assumes that the bandit is risk neutral. It is straightforward to show in the same framework that when the bandit is risk averse, the optimal contract will be a mix of input and output taxes/payments. The next proof considers sheltering activities.

Proof that the optimal contract is a mix of input and output taxes when inputs and outputs can be hidden. Consider the dual problem. To avoid dealing with interactions between sheltering and labor supply, I focus on a problem with infinitely elastic labor supply. This allows me to extract the role of sheltering activities, while leaving the qualitative conclusions unchanged. The proof exploits the stationary bandit’s revenue neutral change in \( \tau_i \) and \( t_i \) around the optimal choice of \( e_i \) and \( H_i \) for the household. The stationary bandit’s revenue neutral infinitesimal change in taxes in sector \( i \) for a given \( e_i \) is characterized by:

\[
\frac{dR}{d\tau_i} = p_i (\alpha_i e_i - H_i) - p_i \frac{\partial H_i}{\partial \tau_i} + e_i dt_i
\]

Total differentiation of the household’s objective function yields:

\[
\frac{dEU}{d\tau_i} = \frac{\partial EU}{\partial \tau_i} d\tau_i + \frac{\partial EU}{dt_i} dt_i
\]

substituting for \( dt_i \) from the last equation:

\[
= e_i p_i d\tau_i \left[ -u(\tilde{\alpha}_i - \overline{\alpha}_i) \right] - e_i p_i \frac{dH_i}{d\tau_i} Eu'd\tau_i
\]

It is immediate to see that \( dEU > 0 \) if \( \tau_i = 0 \) which implies that conditional on a tax on inputs being optimal, it must be that there are also taxes on outputs in the case of uncertainty.

I have already shown that \( dEU > 0 \) if \( \tau_i \). Let’s now show that \( dEU < 0 \) at \( \tau_i = 1 \).
\[ \frac{dE}{d\tau_i} = p_i d\tau_i \left[ -\text{cov}(u', \tilde{\alpha}_i e_i) - \frac{\partial H_i}{\partial \tau_i} E u' \right] \]

\[ = -p_i d\tau_i \frac{\partial H_i}{\partial \tau_i} E u' \]

\[ = -p_i E u' \frac{\partial H_i}{\partial \tau_i} < 0 \]

The first equality follows from the fact that in an interior solution, when \( \tau_i = 1 \), \( \text{cov}(u'(.), \tilde{\alpha}_i) = 0 \) (since at interior solutions for the hiding choice, \( \tilde{\alpha}_i \) where \( H_i = \tilde{\alpha}_i e_i \) are of measure zero.)

The maximum \( \tau_i \) therefore must be interior. A necessary condition for the maximum is therefore \( dE U = 0 \), which obtains by setting the expression for \( dE U = 0 \), ie, whenever:

\[ \tau_i^* = -\frac{\text{cov}(u', e_i \tilde{\alpha}_i)}{\frac{\partial H_i}{\partial \tau_i} E u'} \]

Derivation of the closed form formula for input and input taxes when both can be hidden. The proof of the optimal tax with imperfectly observable inputs and outputs exploits a revenue neutral change in \( t_i \) and \( \tau_i \) and selects amount the sect of revenue constant combinations the one that maximizes household welfare. As explained above, I focus on the problem, given \( e_i \). The bandit’s revenue in sector \( i \) can be written as:

\[ R_i = \tau_i p_i (\alpha_i e_i - H_i) + t_i (e_i - e_i^H) \]

Total differentiation of the bandit’s revenue in sector \( i \) gives:

\[ dR = d\tau_i \left[ p_i (\alpha_i e_i - H_i) - \tau_i p_i \frac{\partial H_i}{\partial \tau_i} \right] + dt_i \left[ e_i - e_i^H - t_i \frac{\partial e_i^H}{\partial t_i} \right] \]

The infinitesimal change in \( t_i \) and \( \tau_i \) is revenue neutral, given \( e_i \), by setting \( dR_i = 0 \). Total differentiation of the household’s utility given \( e_i \) gives:

\[ dU = p_i (-\alpha_i e_i + H_i) d\tau_i - e_i dt_i \]

Expressing \( dt_i \) by the equivalent revenue neutral \( d\tau_i \) from the previous equation gives:

\[ dU = \frac{p_i d\tau_i}{e_i - e_i^H - t_i \frac{\partial e_i^H}{\partial t_i}} \left[ t_i (\alpha_i e_i - H_i) \frac{\partial e_i^H}{\partial t_i} - \tau_i \frac{\partial H_i}{\partial \tau_i} \right] \]

Setting \( dU = 0 \) gives the revenue neutral combination that minimizes distortions:

\[ \tau_i^* = t_i^* (\alpha_i e_i - H_i) \frac{\partial e_i^H}{\partial \tau_i} \]

\[ = t_i^* (\alpha_i e_i - H_i) \frac{\partial e_i^H}{\partial \tau_i} \]

Proceeding similarly, it is straightforward to derive the set of contracts as a function of the distri-
bution of risk aversion in this population and the observability of inputs and outputs. Table 15 characterizes the optimal contracts.

Table 15: Optimal contracts

<table>
<thead>
<tr>
<th>Sheltering activities:</th>
<th>$\tau = 1$</th>
<th>$0 &lt; \tau &lt; 1$</th>
<th>$0 \leq \tau \leq 1$</th>
<th>$\tau = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>$t &lt; 0$</td>
<td>$0 &lt; t &lt; 1$</td>
<td>$0 \leq t \leq 1$</td>
<td>$0 &lt; t &lt; 1$</td>
</tr>
<tr>
<td>Outputs</td>
<td>$0 &lt; \tau &lt; 1$</td>
<td>$0 &lt; t &lt; 1$</td>
<td>$\tau = 0$</td>
<td>$\tau = 0$</td>
</tr>
<tr>
<td>Inputs</td>
<td>$\tau = 1$</td>
<td>$0 &lt; \tau &lt; 1$</td>
<td>$0 &lt; t &lt; 1$</td>
<td>$t = 0$</td>
</tr>
<tr>
<td>Inputs and outputs</td>
<td>$0 &lt; \tau &lt; 1$</td>
<td>$0 &lt; t &lt; 1$</td>
<td>$0 &lt; \tau &lt; 1$</td>
<td>$0 &lt; t &lt; 1$</td>
</tr>
<tr>
<td>Bandits</td>
<td>Risk neutral</td>
<td>Risk averse</td>
<td>Risk neutral</td>
<td>Risk averse</td>
</tr>
<tr>
<td>Households</td>
<td>Risk averse</td>
<td>Risk neutral</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A.5 Time horizon

Let $V(p, F, a_0, M, L, \cdot)$ be the stationary bandit’s indirect utility function. The bandit chooses to be stationary if:

$$\tilde{\pi}V(p, F, a_0, M, L, \cdot) - F > V_r$$

where $\tilde{\pi} = \pi \beta$. Clearly the left hand side is monotonically increasing in $\tilde{\pi}$ and $\exists \pi$ such that:

$\forall \tilde{\pi} \geq \pi$, the bandit chooses to be stationary bandit; $\forall \tilde{\pi} < \pi$, the bandit chooses to be roving bandit.

Factors that reduce the stationary bandit’s expected ability to collect taxes will reduce the return to choose stationary bandit as opposed to the alternative of roving bandit.

The role of the time horizon can also be elicited writing the problem explicitly as an inter-temporal problem. Let $V(p, F, a_0, M, L, \cdot)$ be the stationary bandit’s value function. It is the stationary bandit’s present discounted value when he is stationary. Let the contemporaneous tax revenue be $R$. Let’s suppose that taxes today affect production tomorrow. This is reasonable, for instance, if households’ production and assets are complementary, or if households may migrate out of the village in response to high taxes, and a migration response is slow. Let this state variable be $n_t$, where $n_{t+1} = n(n_t, \tau_t)$ and $\frac{\partial n(n_t, \tau_t)}{\partial \tau_t} < 0$. The stationary bandit’s problem is:

$$V(n_0, p, F, a_0, L, \cdot) = \max_{\{\tau_t\}} \sum_{t=0}^{\infty} \tilde{\pi}^t R(\tau_t, n_t)$$
By the principle of optimality, this can be rewritten as:

\[ V(n_0, p, F, a_0, L, .) = \max_{\{n_0\}} \{R(\tau_0, n_0)\} + \tilde{\pi} V(n_1, p, F, a_0, L, .) \]
\[ = \max_{\{\tau\}} \{R(\tau, n)\} + \tilde{\pi} V(n, \tau, p, F, a_0, L, .) \]

From the first order condition and the envelope theorem:

\[ \frac{\partial R(\tau_t, n_t)}{\partial \tau_t} = \tilde{\pi} \frac{\partial R(\tau_{t+1}, n_{t+1})}{\partial \tau_t} \lambda \]

where

\[ \lambda = - \frac{\partial n (n_t, \tau_t)}{\partial \tau_t} > 0 \]

Clearly, when \( \tilde{\pi} = 0 \), \( \frac{\partial R(\tau^{\tilde{\pi}=0}, n_t)}{\partial \tau_t} = 0 \), which determines the optimal tax in the static problem. However, if \( \tilde{\pi} > 0 \), \( \frac{\partial R(\tau^{\tilde{\pi}=\cdot}, n_t)}{\partial \tau_t} > 0 \). When \( R(\tau_t, n_t) \) is concave, this translates to:

\[ \tau^\tilde{\pi} < \tau^{\tilde{\pi}=0} \]

Furthermore, it is straightforward to show that \( \frac{\partial {\tau^*}(\tilde{\pi})}{\partial \tilde{\pi}} < 0 \) and \( \frac{\partial V(n_0, p, F, a_0, L, .)}{\partial \tilde{\pi}} > 0 \).

### A.6 Economic effects

This discussion focuses on a village with one sector only. The theoretical result remains when introducing various sectors. Under the assumption that without stationary bandit, households are expropriated with probability one, trivially \( e^*_R = 0 \) and \( y^*_R = 0 \). Let \( c(e) = \frac{e^2}{2} \) and the utility function be linear in consumption. Labor supply in this form is isoelastic. Suppose furthermore, that output cannot be hidden, \( k = +\infty \). The households’ labor supply under a stationary bandit in this case is: \( e^*_{SB} = \frac{pa}{2} \) and production \( y^*_{SB} = \frac{pa^2}{2} \). Since production is zero when there is no protection by a stationary bandit, this is also the gain in production from going from no stationary bandit to stationary bandit present.

Of course, that the village will be expropriated with probability 1 in the absence of a stationary bandit is an assumption. In reality, villages without stationary bandits may not be expropriated with probability 1, and villages with a stationary bandit may be expropriated by a successful and uncoordinated roving bandits. Suppose, now, that a village without a stationary bandit may be expropriated by roving bandits with probability \( q \). Labor supply in that case will be \( e^*_R = (1-q)pa \). It now follows that output with a stationary bandit is larger than output without a stationary bandit if and only if \( q > \frac{1}{2} \). Introducing risk aversion reinforces the result that a stationary bandit has a positive effect on economic activity (the lower threshold for \( q \) is lower), since by having the monopoly of violence, he will reduce uncertainty over the level of expropriation. With sufficient prudence, households will respond to the reduction in uncertainty by increasing labor supply.
B Data Appendix

B.1 Validation of the coltan shock

Figure 19: The coltan price shock from satellite

Notes: This figure shows two images of the survey area from satellite at night. The left figure shows the average cloud free lights captured by NASA-NOAA satellites from the survey area in 1999. The right figure is the equivalent figure for 2000. The orange lines indicate international borders. On the left is the Democratic Republic of Congo. On the right, and from North to South: Uganda, Rwanda, Burundi, Tanzania. The endowments of the villages are as indicated in the figure legend. In the year 2000, a very large town emerges in the northern part of the picture. This town is close to Goma, the provincial capital of North Kivu, which is the richest region in coltan. As the price of coltan skyrocketed, miners, traders and individuals whose activities related to mining had an income shock, and consumption increased in the main towns in the proximity of the provincial capitals.
Figure 20: coltan shock and behavioral responses: number of marriages and occupational choice

Notes: This figure plots the number of recorder marriages on years. The left vertical axis indicates the average number of marriages and the horizontal axis indicates the year. The solid line indicates the average number of marriages in coltan villages, by year. The dashed line indicates the average number of marriages in villages not endowed with coltan, by year.
## B.2 Validation of the data

Table 16: Survey and ACLED battles

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conquest attempts, survey</td>
<td>0.0138***</td>
<td>0.0298***</td>
<td>0.0565***</td>
<td>0.0679***</td>
<td>0.100***</td>
<td>0.0872***</td>
<td>0.122***</td>
<td>0.102***</td>
</tr>
<tr>
<td></td>
<td>(0.00396)</td>
<td>(0.00674)</td>
<td>(0.0107)</td>
<td>(0.0167)</td>
<td>(0.0233)</td>
<td>(0.0268)</td>
<td>(0.0291)</td>
<td>(0.0359)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.00718*</td>
<td>0.0295***</td>
<td>0.0515***</td>
<td>0.0817***</td>
<td>0.126***</td>
<td>0.203***</td>
<td>0.248***</td>
<td>0.614***</td>
</tr>
<tr>
<td></td>
<td>(0.00373)</td>
<td>(0.00636)</td>
<td>(0.0101)</td>
<td>(0.0158)</td>
<td>(0.0220)</td>
<td>(0.0253)</td>
<td>(0.0274)</td>
<td>(0.0339)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,102</td>
<td>2,102</td>
<td>2,102</td>
<td>2,102</td>
<td>2,102</td>
<td>2,102</td>
<td>2,102</td>
<td>2,102</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.107</td>
<td>0.133</td>
<td>0.146</td>
<td>0.172</td>
<td>0.232</td>
<td>0.271</td>
<td>0.281</td>
<td>0.396</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1. This table presents the results on the OLS regression of battles recorded by ACLED on battles recorded by my survey. To assign battles recorded by ACLED to the survey villages, I derived circles of varying circumferences to each of my survey villages. For each circle, I computed the number of geo-located ACLED battles that fall inside the circle. Thus, for each village, the assigned number of ACLED battles is the number that is closer than the radius of the selected circle. Across columns, I present the results for circles of the following radius: 1km (column 1), 2km (column 2), 5km (column 3), 10km (column 4), 15km (column 5), 20km (column 6), 25km (column 7), 50km (column 8). Standard errors are clustered at the village*year level.
Figure 21: Violence compared with known historical rebellions

Notes: This figure plots the number of attacks on the sample villages by different armed groups recorded in my survey, by year and compares them to well known dates for known historical rebellions. The left axis indicates the number of attacks recorded in the sample, and the horizontal axis indicates the year. The blue line indicates attacks by the AFDL as recorded in the survey. The black line indicates attacks by the RCD as recorded in the survey. The red, dashed line indicates the number of attacks by the CNDP group as recorded in the survey. Correspondingly, the vertical blue lines at 1996 and 1997 mark the well known dates for the launch and end of the AFDL rebellion as known from historical accounts; the vertical dotted lines at 1998 and 2003 mark the well known dates for the launch and end of the RCD rebellion as known from historical accounts; the vertical dashed line at 2004 mark the well known date for the launch the CNDP rebellion as known from historical accounts. The dates of the attacks recorded from the survey coincide exactly with well known historical rebellions. The source for this graph is the attacks dataset, which contains less, more detailed information on attacks.
Figure 22: Survey data compared to ACLED data

Notes: This figure plots the total number of attacks on the sample villages recorded in my survey and recorded by ACLED in the neighborhood of my survey villages, by year. The solid line represents the total number of attacks recorded in my survey, and the dotted and dashed lines represent the number of attacks recorded from ACLED, in the neighborhood of my survey villages. To assign attacks recorded by ACLED to the survey villages, I derived circles of varying circumferences to each of my survey villages. For each circle, I computed the number of geo-located ACLED violent events that fall inside the circle. Thus, for each village, the assigned number of ACLED battles is the number that is closer than the radius of the selected circle. The dotted line indicates the resulting number of violent events in a circle of 10km radius. The dashed and dotted line, 5 km radius. The dashed line, 2 km radius. The solid line captures the well-known phases of the Congo Conflict. The number of attacks rises in 1998 drastically, with the Second Congo War, an in 2000, with the coltan shock. It then decreases with the post-conflict period, and rises again in 2009, 2010. This last rise is the rise in attacks by the FDLR as a result of the Kimia II operation that led them to pillage. The ACLED dataset does not capture these trends. This provides additional confidence in my survey attacks data. The source I used for villages in my survey for this graph is the village dataset, which contains more, less detailed information on attacks.