The quality of political institutions and the curse of natural resources

Antonio Cabrales† and Esther Hauk‡

February 2009

Abstract

We propose a theoretical model to explain empirical regularities related to the curse of natural resources. This is an explicitly political model which emphasizes the behavior and incentives of politicians. We extend the standard voting model to give voters political control beyond the elections. This gives rise to a new restriction into our political economy model: policies should not give rise to a revolution. Our model clarifies when resource discoveries might lead to revolutions, namely, in countries with weak institutions. Natural resources may be bad for democracy by harming political turnover. Our model also suggests a non-linear dependence of human capital on natural resources. For bad political institutions human capital depends negatively on natural resources, while for high institutional quality the dependence is reversed. This theoretical finding is corroborated in cross section regressions.

JEL Codes: D72, H52, O13.

Keywords: Curse of natural resources, quality of political institutions, revolution, human capital.

* A previous version of the paper was circulated under the title: Democracy and the curse of natural resources. We thank Pablo Fleiss and Ognjen Obucina for valuable research assistance. We are especially indebted to Albert Marcet who helped us derive the asymptotic distribution of our estimator. We also thank Facundo Albornoz-Crespo, Matteo Cervellati, José García-Montalvo, Eleonora Pattacchini and participants in seminars at the University of Zurich, the Institute of Advanced Studies in Vienna, the University of Alicante, Cambridge University, the University of Brescia for helpful comments. We gratefully acknowledge financial support from the Spanish Ministry of Science and Technology under grants CONSOLIDER-INGENIO 2010 (CSD2006-0016), SEJ2006-01717 and SEJ2006-11665-C02-00.

† Universidad Carlos III de Madrid, e-mail: antonio.cabrales@uc3m.es
‡ Institut d’Anàlisi Econòmica CSIC, e-mail: esther.hauk@iae.csic.es
1 Introduction

Until World War II the economic profession tended to believe that natural resources were an unqualified blessing for the nation owning them. However, in the post world-war II period the evidence against this belief started accumulating: many resource rich countries grew very slowly and economists started to talk about the curse of natural resources. There is a large number of empirical papers which find evidence of this curse (e.g. Sachs and Warner (1995, 1997, 1999, 2001), Mehlum et al. (2006), Gylafson (2004), Strauss (2000)). Some authors (Sala-i-Martin (1997) and Doppelhofer et al. (2000)) have even classified natural resources as one of the ten most robust variables with a significantly negative effect on growth in empirical studies.

To summarize, there seems to be an empirical consensus on the following:

Fact 1 The curse of natural resources: countries rich in natural resources grow slower on average than natural resource poor countries.

However, there are many important outliers. Some resource rich countries have grown very fast (e.g. Botswana, Canada, Australia, Norway) while others have grown very slowly (e.g. Nigeria, Zambia, Sierra Leone, Angola, Saudi Arabia, Venezuela). It seems fair to claim that:

Fact 2 The cross-country evidence is inconsistent with a monotonic effect of resources on development/growth: (Robinson et al. (2006))

We therefore need to understand when natural resources are a blessing and when they are a curse. The empirical literature has taken a step in this direction and it defines policy failure as the prime cause of the underperformance of resource rich countries. It also points to a reason why these policy failures occur. Namely:

1 Acemoglu et al. (2003) show that Botswana has the highest per capita growth of any country in the world in the last 35 years. The natural resources of Botswana are diamonds. This country had very bad starting conditions for growth (extremely low education levels, bad infrastructure, etc) but “good” institutions.

2 Some countries which have been fairly rich in resources in 1970 that grew rapidly in the next 20 years are Malaysia, Mauritius and Iceland (see Sachs and Warner (2001)). Gylafson (2001a) additionally lists Indonesia and Thailand as countries attaining both long-term investment exceeding 25% GDP and per capita GNP growth exceeding 4% per year on average from 1970 to 1998. Also the so-called Scandinavian catch-up in the late nineteenth century was based on natural resources.
**Fact 3** The quality of institutions is decisive in determining whether natural resources are a blessing or a curse.³

Institutions are linked to the behavior of politicians, as they limit their discretion and define the policy space. The quality of institutions is also indicative of the level of democracy of a country. More democratic countries tend to have better institutions and are therefore less likely to be cursed by natural resources. But empirical findings also suggest a reverse causality known as the political Dutch disease.⁴

**Fact 4** Natural resources have antidemocratic properties: oil and mineral wealth tends to make states less democratic (Ross, 2001, Lam and Wantchekon, 2002, Jensen and Wantchekon, 2004, Damania and Bulte, 2008).

Moreover, in countries with weak institutions natural resources are one of the main sources of civil war and revolution.

**Fact 5** Many revolutions are linked to rents derived from natural resources (Collier and Hoeffler, 1998). In particular, oil, gemstones, minerals and other lootable resources are associated with civil conflict while agriculture is not.⁵

The theoretical contribution of the present paper is threefold:

1. We propose the first theoretical model that incorporates and explains the five empirical facts outlined above.

³Mehlum, Moene and Torvik (2006) show that the effect of resources on growth is positive (negative) when institutions are good (bad) using Sachs and Warner’s (1995) data. The same paper as well as Boschini, Petterson and Roine (2003) show that the direct negative effect is stronger for minerals than other resources and institutions are more decisive for the effect of minerals than other resources.

⁴The usual argument explaining why natural resources harm democratization is based on the incumbent’s discretion over the distribution of natural resources. A noticeable exception is Morrison (2007) who argues that even in an ideal scenario where natural resources are funneled away from nondemocratic governments toward the citizens, natural resources would still hinder democratization. His model is based on Acemoglu and Robinson’s (2006) theory of democratization in which the distributional struggle between the poor and the rich is a reason for democratization. Morrison shows that natural resources reduce the need for redistribution by the rich: if the natural resource revenue is high enough, the poor may no longer prefer a positive tax rate.

⁵For an overview on the empirical literature on the link between civil unrest and natural resources see Ross (2004).
2. We present an *explicitly political model* which emphasizes the behavior and incentives of politicians. This is key, since there is a clear understanding that the behavior of government/politicians is fundamental to explain the economic performance in resource abundant countries (Newberry (1986, p.334)).

3. We extend the standard voting model to give voters political control beyond the elections. Democratic institutions are often imperfect, and electoral competition could be weak. But in our model, as in reality, citizens have instruments in addition to elections that allow them to avoid policies which could cause them big welfare losses. We introduce these considerations in the model by assuming that citizens can initiate a revolution.\(^6\) This gives rise to a new restriction into our political economy model: policies should not give rise to a revolution. We will refer to this new constraint as the *no-revolution constraint*.

The existing theoretical literature concentrates mainly on explaining the “curse” (Fact 1),\(^7\)\(^8\) and suggests ways to avoid the curse.\(^9\) This line of research ignores the role of government and therefore cannot explain why governments do not choose the good policies in the first place.\(^10\) We need explicitly political models to understand when natural resources are a blessing and when they are a curse.

To our knowledge the first explicitly political model in this area was developed by Robinson et al. (2006). Their model explains empirical facts 2 and 3. In their paper there are two periods, with elections at the end of the first period. In the first period, natural resources are discovered. The

---

\(^6\)It need not be violent, although we will assume it causes some economic disruption. General strikes are an example of voters’ control beyond the elections.

\(^7\)For a list of explanations for the natural resource trap and their empirical support see Strauss (2000).

\(^8\)There is no generally accepted explanation for the curse so far. The one with maybe most empirical support is the “Dutch disease” explanation which goes as follows: the discovery and exploitations of natural resources like oil typically leads to large profits. These profits encourage entry into the industry at the expense of other sectors, expand national income and increase demand with a resulting inflationary pressure. At the same time more foreign currency enters the country which appreciates the real exchange rate. Export profits in the non-boom sector fall sharply which attracts even more capacity into the boom sector. The long-run results once the boom is over are stagflation and an over-valued real exchange rate.

The Dutch disease is preventable by good policies; e.g. Indonesia avoided the disease after its oil discovery by consistently devaluating its currency.

\(^9\)See e.g. Birdsall et al. (2000).

\(^10\)Rent-seeking and corruption are explanations that have been put forward. In these models the state is an aggregator of pressure from interest groups (Becker-Olson approach) which as Robinson et al. (2006) pointed out ignores incentives of politicians who often have a large amount of autonomy from interest groups.
incumbent government has to decide which proportion of the resources to extract then, and how much to leave for the following period. The government can consume the resource income, or use it to influence election outcomes by offering employment in the public sector, which is relatively inefficient. The main result of the paper is that politicians tend to overextract resources in the first period because they only care about the future resources if they remain in power. Moreover, the public sector will be inefficiently large. Institutions are decisive for the overall impact of resource booms because they determine the extent to which political incentives can really influence policy outcomes.

While the size of the public sector and the extraction path of natural resources are clearly relevant issues, there are other important channels from natural resources to growth that are unexplored by Robinson et al. (2006). In particular, human capital accumulation or education. One danger of natural resources (Gylfason (2001a)) is the neglect of education, since the country can live well over an extended period even with a weak commitment to education. But since we know that increased education is conducive to higher growth levels (Barro (2001), Barro and Lee (2001), Gylfason and Zoega (2004)), this reduced commitment to education will surely cost those countries in terms of long-run growth. For this reason, it is difficult to explain the higher persistence of growth in resource-rich Scandinavia than in Latin America (especially resource-rich countries such as Argentina and Chile) without remarking on the educational gap that emerged between the two groups of countries over the period 1870-1910 and which remained large throughout the twentieth century (see Bravo-Ortega et al. (2002)).

In this paper we build an explicitly political model to explain when natural resource discoveries lead to higher or lower education levels. We are only interested in publicly owned resources,\textsuperscript{11}

\textsuperscript{11}Bulte and Damania (2008) present an explicitly political model in which resources are not publicly owned. In the resource sector entrepreneurs claim a fraction of the resource stock and extract from their private substock. In their model entrepreneurs decide whether to enter resource extraction which has diminishing returns or manufacturing which has increasing returns. Hence moving into manufacturing yields external benefits. Production in both sectors requires a sector specific semi public good provided by a purely self-interested government. To get this good the sectors offer payments/bribes to the government, who decides how much of the good to produce in return. The government might be challenged by a political opponent. The manufacturing sector bribes too little and hence gets too little of the sector-specific good, since firms do not internalize the spill-over benefits from production. As in our paper, the stronger the potential rival, the more the incumbent government has to take welfare maximizing considerations into account to remain in power. As a result the resource curse emerges if there is no credible opposition or political transaction costs are high. With strong political competition the government cannot strain far from the income maximizing path and hence resource booms are not detrimental for growth.
such as oil, gas and minerals.\textsuperscript{12} Politicians are purely self-interested and would like to consume the resource wealth themselves, but political pressure obliges them to redistribute at least a part of it to voters. This redistribution can take the form of: (i) a direct transfer or (ii) a subsidy for the investment in human capital, which has a positive spillover on the entire population. The incumbent government faces political pressure from two sources: an election and the possibility of a revolution. We model political opposition by the existence of a competitive fringe. Additionally, we assume that political transitions are not without costs. These costs depend on the quality and transparency of political institutions and the level of human capital of the fringe players. The efficiency of the fringe increases in human capital. For low levels of education the fringe will always be less efficient than the incumbent at managing natural resources. Whether or not this situation can be reversed for highly educated fringe players crucially depends on the quality of political institutions. Hence, the function characterizing the efficiency of the fringe players gives us a measure for institutional quality.\textsuperscript{13}

Besides the political competition there is always a possibility of a revolution. If the revolution is successful, natural resources fall into the hand of the voters who divide the gains equally among themselves. These gains now depend on the management skills of voters. We assume that voters are better at managing natural resources the higher their level of education.

We establish the following main results:

1. If the fringe wins the election, human capital increases with the size of the stock of natural resources.

2. If the government wins the election, human capital is a non-increasing function of natural resources.

3. If the government does not have to worry about revolution, human capital is constant.

4. If revolution is a binding constraint, human capital decreases in natural resources.

5. Revolution is less likely to be a threat, the better are a country’s political institution.

\textsuperscript{12}In all petrostates the government maintains explicit legal ownership of below-ground reserves irrespective of surface property rights (see footnote 12 in Lam and Wantchekon (2002)). Most OPEC governments put the resources under national control in particular in the 60s and 70s.

\textsuperscript{13}This assumption will be justified at length in the model section.
6. The probability that the incumbent is reelected may increase with natural resources and this is more likely for countries with bad institutions.

These results confirm that our explicitly political model captures the five empirical facts mentioned above. Our model clarifies when resource discoveries might lead to revolutions (Fact 5), namely, in countries with weak institutions. In our model, natural resources may be bad for democracy because they can harm political turnover (Fact 4). Our model suggests a non-linear dependence of human capital on natural resources (Fact 2). For low levels of institutional quality human capital depends negatively on natural resources, while for high levels of institutional quality the dependence is reversed (Fact 3). Since natural resources are bad when the government wins the election, and this probability may increase with natural resources, especially in countries with bad institutions, natural resources are a curse on average (Fact 1).

Empirical facts 1 to 3 were stated in terms of growth. We do not model growth directly, but use human capital/education instead, which is an established engine of growth (Barro 2001). An explicitly dynamic model would allow to study, for example, the dynamics of capital accumulation, at the expense of a considerable complication in its exposition and development. Our model allows to explain the empirical facts already discussed as arising from the effects of education on growth. In addition, our model yields direct predictions for the effect of natural resources on education, which can then be tested empirically. Existing empirical studies report conflicting results of the effect of natural resources on education.14 The most complete study is the one by Stijns (2004), who discusses the different indicators used for resource abundance and human capital accumulation and shows that the conclusion on the link between these two is sensitive to the indicators chosen. Simple correlation coefficients and regressions switch from positive to negative depending on which resource abundance and which human capital indicator is used. This evidence might be consistent with the non-linear dependence of human capital on natural resources predicted in our model. This, however, has to be tested by including a variable that measures the quality of the political institutions of a country in the regressions. There does not exist such an empirical study. We therefore run our own regressions which confirm that the non-linear relationship is indeed driven by the institutional quality variable. Since OLS regressions might suffer from an endogeneity problem we propose a two-stage non-linear least square estimator in which the institutional quality index

14Gylfason (2001a, 2004) establishes an inverse relationship of human capital measured as public expenditure on education, expected years of schooling for girls and gross secondary-school enrollment with the share of natural capital in national wealth. However, the results seem to be driven by very few countries.
is instrumented by the average institutional quality index of the neighboring countries. In the Appendix we show that this estimator is consistent and derive its asymptotic distribution.

The remainder of the paper is organized as follows. Section 2 presents the model and solves it. Section 3 test the model empirically. Section 4 concludes.

2 The Model

Assume a country owns a stock of natural resources whose rents generate a discounted present value $W$. These resources are publicly owned and will therefore be managed by politicians. Politicians are motivated solely by self-interest, hence the government would like to keep the gains from the resources for itself, but it will only be able to benefit from the resource discovery if it remains in power. There are two potential threats for the government’s power: an upcoming election and the possibility of a revolution. Before the election the different political parties propose a contract to voters. The contract consists of a direct money transfer to voters and a per unit subsidy for human capital accumulation. Then the election takes place. Once the election outcome is known, investment in human capital is made. The contract proposed by the winning party will be implemented unless voters decide to make a revolution and the revolution is successful. A successful revolution means that citizens grab the natural resources and split them equally among themselves while all productive activity is forgone. We now describe the different steps in detail. We will start with the electoral process.

In the elections, the government $G$ faces the opposition of a competitive fringe. In other words, the opposition consists of several parties that compete among themselves. The unique policy issue is how to distribute the rents generated by the natural resource. We assume that the value of resources depends on the winner of the elections: its value will be $W$ if managed by the incumbent government and $\delta(e)W$ if managed by one of the fringe parties, where $e$ stands for human capital and $\delta(e) \geq 0$ for all $e$. Furthermore, we assume that $\delta(0) = \bar{\delta} < 1$ and $\delta'(e) > 0$, i.e. for low levels of education the fringe is always less efficient than the government at managing natural resources but the competence of the fringe increases with human capital. The underlying idea of this assumption is that political transitions are not without costs. Usually, the incumbent party obtains certain advantages from being in government. For example, the whole apparatus of the state can be used by this party to get access to information and other resources. In addition, the

\[\text{See, e.g. Cox and Katz (1996), who document empirically the sources of incumbency advantage.}\]
incumbent politicians may become more able over time by a simple learning by doing process. The size of the incumbency advantage depends on the quality of the political institutions. In some countries, basic institutions work independently of who is in office, while in other places even secretarial jobs depend on the party in power. In the latter case, which is also known as the “spoils system,” a change in government implies new workers in key jobs, which obviously leads to severe inefficiencies. A lot of information has to be rediscovered, many things have to be learned again: how costly this change is will depend on the human capital of the new workers and on the transparency of institutions. It can be easier or harder to hide existing information to the newcomers in power. Our function $\delta(e)$ measures this efficiency loss and is therefore a proxy for the goodness of a country’s political institutions. The better are the institutions of a particular country, the better is its political competitive fringe at managing natural resources, which for

16 Padró-i-Miquel and Snyder (2006) demonstrate that legislators’ “Effectiveness never declines with tenure, even out to nine terms. The increase in effectiveness is not simply due to electoral attrition and selective retirement, but appears to be due to learning-by-doing.”

17 Jonas and Jones (1956) cite arguments from the earliest study of turnover of state personnel by Professor Martin L. Faust against the spoils system.

“The spoils system entails heavy turnover in personnel which periodically results in the scrapping of all or nearly all accumulated experience. It places inexperienced and incompetent persons in responsible administrative positions. Since it is predicated upon rewards and favors, it introduces favoritism and partiality in the conduct of the public business and limits the access to the public service of young people of capacity and promise. The spoils system renders impossible continuity in administrative policy and destroys morale within the service. It ‘leads to oligarchy and autocracy by helping bosses get control of the party machinery.’ Moreover, the prevalence of the spoils system in state government makes difficult effective federal-state co-operation and at the same time encourages the growth of bureaucracy at both levels.” (p.755)

18 Notice that we do not exclude the possibility that the fringe might become more efficient than the government at managing natural resources: if political institutions are good $\delta(1) > 1$ but for countries with bad political institutions there will always be an efficiency loss. In this countries $\delta$ will be very bounded and low, i.e. $\delta(1) << 1$.

19 An alternative and also interesting interpretation of the $\delta(e)$ function could be that it somehow captures how costly it is for the opposition to get access to elections. In countries with worse institutions the costs for the opposition to get access to elections are higher and it will use part of the resource wealth to recover those costs once in power.

20 To fully understand our assumptions it will be useful to comment on what would happen in a dynamic extension of the model after a change in government. If the opposition wins, today’s incumbent becomes part of tomorrow’s opposition and today’s fringe becomes the new incumbent. We argue that we can use the same assumptions about relative efficiency of (new) fringe and incumbent as in the static model. The old fringe is now an incumbent and has gathered experience on making the institutions function with his team. He has privileged information and the incumbency advantage. The former incumbent (the new fringe player) does not keep his former efficiency advantage
simplicity is the only task of politicians in our model. While our argument is more general (and it should be thought of in these general terms), the quality of institutions also affects the resource sector directly. In some countries, this sector is fairly independent of the incumbent government, because resource extraction is handled by privately-owned multinational corporations, but there are other countries in which the firms are state-owned and employment in these firms might be subject to changes in government.\footnote{PeMex (Petroleos Mexico) is an example of a state owned firm in which employment depends on the party in power: Arellano Gault and Klinger (2004) refer to PeMEX as a politically sensitive agency.}

There are two ways to transfer resource rents to voters, (i) via a direct (per capita) transfer $w$ and (ii) via a per unit subsidy $\pi$ for the investment in human capital. The individual’s level of human capital $e$ together with the average level of human capital $\bar{e}$ determines each individual’s marginal productivity (salary) $\omega$ in the following way (where we assume $\alpha + \beta < 1$):

$$\omega = ke^\alpha e^\beta$$

Hence there is a positive externality (spillover) for society as a whole if an individual invests in human capital. We assume that the monetary cost of acquiring a unit of human capital is $\lambda$. Given the promised transfers, the voter decides on his own level of human capital by maximizing his utility. Hence, the program of the voter is:

$$\max_e U(w + ke^\alpha e^\beta - (\lambda - \pi)e)$$

The first order conditions of this (concave) problem give $\lambda - \pi = \alpha ke^{\alpha - 1} e^\beta$. Since all voters are identical we can assume that in equilibrium $e = \bar{e}$. Therefore

$$\lambda - \pi = \alpha ke^{\alpha + \beta - 1}. \quad (1)$$

We will refer to equation (1) as the voter’s incentive compatibility constraint: it tells us the level of human capital of a voter given the size of the subsidy $\pi$. Using this constraint, we can talk directly about the level of human capital $e$ resulting from the transfers instead of discussing the size of the subsidy $\pi$. Hence voter’s material utility can be rewritten as a function of the direct transfer $w$ and the level of human capital $e$, namely

$$U(w + ke^{\alpha + \beta} - \alpha ke^{\alpha + \beta}). \quad (2)$$
There is a continuum of voters with total mass $n$. Voters care about the promised utility by the competing parties but also have some ideological concerns. The fringe parties are perceived by voters as ideologically equivalent, hence we can assume that the equilibrium behavior of fringe players will be identical (we focus on a symmetric equilibrium). From now on, all endogenous variables will be indexed by the political actor offering them. Thus, we have $\omega_i, w_i, e_i, \bar{e}_i, \pi_i$ with $i \in \{G,F\}$ where $G$ stands for "Government" and $F$ for "Fringe".

The electoral process is a version of the probabilistic voting model and works in the following way:

Voters are located in the interval $[0,1]$. The utility of a voter $v \in [0,1]$ when offered a policy that delivers “material” utility $U_G$ from the government is denoted

$$u(v, U_G) = U_G - \theta v$$

The utility of a voter $v \in [0,1]$ when offered a policy that delivers “material” utility $U_F$ from the competitive fringe is denoted

$$u(v, U_F) = U_F - \theta (1 - v)$$

where $\theta$ denotes the strength of purely ideological concerns.

In addition, in every election there is an unexpected “aggregate shock” $\varepsilon \sim U[-A,A]$ to the utility that shifts preferences of all the voters in favor or against the incumbent. We add this shock to the preferences toward the incumbent.

$$u(v, U_G) + \varepsilon$$

The proportion of voters preferring $G$ over $F$ is then:

$$\min \left\{ \max \left\{ 0, \frac{1}{2} + \frac{U_G - U_F}{2\theta} + \frac{\varepsilon}{2}\theta \right\}, 1 \right\}$$

Thus, the ex ante probability that the incumbent wins the election, given promises $U_F, U_G$ is:

$$\Pr \left[ \min \left\{ \max \left\{ 0, \frac{1}{2} + \frac{U_G - U_F}{2\theta} + \frac{\varepsilon}{2}\theta \right\}, 1 \right\} \geq \frac{1}{2} \right].$$

Hence, the incumbent wins for all $\varepsilon > \varepsilon_1$ where $\varepsilon_1$ makes $\frac{1}{2} + \frac{U_G - U_F}{2\theta} + \frac{\varepsilon_1}{2}\theta = \frac{1}{2}$. Thus, $\varepsilon_1 = -(U_G - U_F)$. The probability of winning for the incumbent is equal to
\[
\Pr [\varepsilon > \varepsilon_1 = -(U_G - U_F)] = \min \left\{ \max \left\{ 0, \frac{A - \varepsilon_1}{2A} \right\}, 1 \right\} \]

(3)

\[
= \min \left\{ \max \left\{ 0, \frac{1}{2} + \frac{U_G - U_F}{2A} \right\}, 1 \right\}.
\]

The incumbent cannot win if \(\frac{1}{2} + \frac{U_G - U_F}{2A} < 0\) which implies that \(A < -(U_G - U_F)\). On the other hand, the incumbent wins with probability 1 for \(A < (U_G - U_F)\).

After the election results, the citizens decide whether or not to make a revolution.\(^{22}\) Hence, if a revolution takes place, this happens after the acquisition of human capital. We assume that a revolution is costly (its marginal cost is \(c\)) and it is successful with probability \(q\). In case of a successful revolution, the citizens manage the natural resources and obtain an equal split of these resources. Just like with the fringe politicians, the natural resources that go to the citizens after the revolution increases with human capital. We model this by assuming that the natural resources that go to citizens after a successful revolution is \(\gamma(e) \frac{W}{n}\) with \(\gamma'(e) > 0\).\(^{23}\) If the revolution fails, the original contract proposed by the winner of the elections is imposed. Before analyzing the model further we summarize the timing of the model.

**Timing of the model**

1. Resource discovery of size \(W\)

2. The incumbent and the fringe opposition offer contracts \((w, \pi)\) to voters.

3. Nature chooses the aggregate preference shock toward the incumbent.

4. Voting takes place.

5. The election outcome becomes known and human capital is acquired.

\(^{22}\)We think of a revolution as a threshold public good problem. At least \(x\) people have to go, or the revolution will not take place. This modeling choice leaves unanswered the question of who does the revolution. In our model there is a natural candidate: the group of voters ideologically most distinct from the winning party.

\(^{23}\)It is very likely that after a revolution citizens who manage the natural resources themselves do not entirely rely on former experts. This is captured by the \(\gamma(e)\) function that parallels the \(\delta(e)\) function. Alternatively, we could assume that the management skills for natural resources of revolutionaries are independent from natural resources, but the more educated they are, the lower the cost of revolution or the higher the probability of success.
6. Citizens decide whether or not to make a revolution.

7. If no revolution is made, the contract proposed by the winning party is implemented. In case of a revolution nature determines whether or not it is successful (probability \( q \)).

- if successful, citizens manage the natural resources themselves and there is no productive activity.
- if the revolution fails, the original contract of the election winner is imposed.

Given these assumptions political parties always want to avoid the revolution. The no-revolution constraint requires that the promised contracts have to be at least as good as the outcome of the revolution, i.e.

\[
U(w + ke^{\alpha+\beta} - ake^{\alpha+\beta}) \geq qU(\gamma(e)W/n) + (1-q)U(w + ke^{\alpha+\beta} - ake^{\alpha+\beta}) - c
\]

which simplifies to:

\[
U(w + ke^{\alpha+\beta} - ake^{\alpha+\beta}) \geq U(\gamma(e)W/n) - \frac{c}{q} \tag{4}
\]

We first observe that:

**Lemma 1** Revolution is a potential threat only to the incumbent government.

**Proof.** First notice that competition among the fringe players drives their profits down to zero. The equilibrium offer by the fringe can thus be obtained by maximizing the consumers’ utility subject to the resource constraint (what we call the fringe program). To see why in equilibrium the fringe does not take the no-revolution constraint into account, suppose that the solution to the above described fringe program (call it program 1) does not satisfy the no-revolution constraint (the only problematic case). Then one could obtain an alternative solution by imposing the constraint (call this the solution to program 2). But the solution to program 2 can only decrease the utility of agents (with respect to the solution of program 1), which can only worsen the constraint, leading to a contradiction. The government, on the other hand, does keep some of the resource rents for itself. Therefore revolution might be a threat for the government. ■

We will now formally state the maximization problems of the fringe players and of the government.
2.1 The fringe problem

Due to competition among fringe players, the fringe maximizes the consumers’ utility subject to the resource constraint \( \delta(e_F)W/n - w_F - \pi_F e_F \geq 0 \). Using the incentive compatibility constraint of voters (1) the resource constraint can be rewritten as \( \delta(e_F)W/n - w_F - \lambda e_F + ak e_F^{\alpha+\beta} \geq 0 \) and we can talk about the fringe choosing \( e_F \) instead of \( \pi_F \). Hence the fringe problem is:\(^{24}\)

\[
\max_{e_F, w_F} U(w_F + ke_F^{\alpha+\beta} - \alpha k e_F^{\alpha+\beta})
\]

subject to \( \delta(e_F)W/n - w_F - \lambda e_F + ak e_F^{\alpha+\beta} \geq 0 \)

Since there is competition among fringe players, the resource constraint has to be satisfied with equality, therefore the fringe problem becomes:

\[
\max_{e_F, w_F} U \left( \delta(e_F)W/n - \lambda e_F + ak e_F^{\alpha+\beta} \right)
\]

which simplifies to

\[
\max_{e_F} U \left( \delta(e_F)W/n - \lambda e_F + k e_F^{\alpha+\beta} \right)
\]

The first order condition is

\[
\delta'(e_F)W/n + (\alpha + \beta)k e_F^{\alpha+\beta-1} - \lambda = 0 \tag{5}
\]

In this way, we have that

\[
\frac{\partial e_F}{\partial W/n} = \frac{-\delta'(e_F)}{(\alpha + \beta)(\alpha + \beta - 1)k e_F^{\alpha+\beta-2} + \delta''(e_F)W/n}
\]

Since we know that if the decision is optimal \((\alpha + \beta)(\alpha + \beta - 1)k e_F^{\alpha+\beta-2} + \delta''(e_F)W/n \leq 0\) (to guarantee the satisfaction of second order conditions), then if \( \delta'(e_F) \geq 0 \), the effect of increasing \( W \) in \( e_F \) is positive. We summarize this observation in:

**Proposition 1** When the fringe wins the election, human capital is positively related to the amount of natural resources.

---

\(^{24}\)Profits are really \( \left( \delta(e_F)W/n - w_F - \lambda e_F + ak e_F^{\alpha+\beta} \right) \min \left\{ \max \left\{ 0, \frac{1}{2} + \frac{W_F-U_G}{2A} \right\}, 1 \right\} \), but notice that \( \min \left\{ \max \left\{ 0, \frac{1}{2} + \frac{W_F-U_G}{2A} \right\}, 1 \right\} \) being a probability is always bigger than zero, thus it never affects whether the constraint is binding or not.
2.2 The government problem

The government maximizes its own utility subject to the no-revolution constraint:

\[
\max_{e_G, w_G} \left( \frac{W}{n} - w_G - \lambda e_G + a k e_G^{\alpha + \beta} \right) \times \min \left\{ \max \left\{ 0, \frac{1}{2} + \frac{U(w_G + k e_G^{\alpha + \beta} - a k e_G^{\alpha + \beta}) - U_F}{2A} \right\}, 1 \right\}
\]

subject to \(U(w_G + k e_G^{\alpha + \beta} - a k e_G^{\alpha + \beta}) \geq U(\gamma(e_G)W/n) - \frac{c}{q}\)

In order to derive some analytical solutions, we further assume that \(U(x) = \ln(x)\). Then

\[
\max_{e_G, w_G} \left( \frac{W}{n} - w_G - \lambda e_G + a k e_G^{\alpha + \beta} \right) \times \min \left\{ \max \left\{ 0, \frac{1}{2} + \frac{\ln(w_G + k e_G^{\alpha + \beta} - a k e_G^{\alpha + \beta}) - U_F}{2A} \right\}, 1 \right\}
\]

subject to \(w_G + k e_G^{\alpha + \beta} - a k e_G^{\alpha + \beta} \geq \gamma(e_G)W/n \exp \left( -\frac{c}{q} \right)\)

We have to distinguish two cases: (i) the no-revolution constraint does not bind at the optimum and (ii) the no-revolution constraint binds at the optimum.

Case (i): unconstrained solution

If the no revolution constraint does not bind (true for sufficiently low values of \(\exp \left( -\frac{c}{q} \right)\)), we have:

\[
\max_{e_G, w_G} \left( \frac{W}{n} - w_G - \lambda e_G + a k e_G^{\alpha + \beta} \right) \times \min \left\{ \max \left\{ 0, \frac{1}{2} + \frac{\ln(w_G + k e_G^{\alpha + \beta} - a k e_G^{\alpha + \beta}) - U_F}{2A} \right\}, 1 \right\}
\]
and the first order conditions are:

\[ 0 = -\left( \frac{1}{2} + \frac{\ln(w_G + ke_G^{\alpha+\beta} - \alpha ke_G^{\alpha+\beta} - U_F)}{2A} \right) + \]

\[ \left( \frac{W/n - w_G - \lambda e_G + \alpha ke_G^{\alpha+\beta}}{2A(w_G + ke_G^{\alpha+\beta} - \alpha ke_G^{\alpha+\beta})} \right) \frac{1}{2A(w_G + ke_G^{\alpha+\beta} - \alpha ke_G^{\alpha+\beta})} \]

\[ 0 = -\lambda + \alpha k(\alpha + \beta) e_G^{\alpha+\beta} - 1 \left( \frac{1}{2} + \frac{\ln(w_G + ke_G^{\alpha+\beta} - \alpha ke_G^{\alpha+\beta} - U_F)}{2A} \right) \]

\[ + \left( \frac{k(1 - \alpha)(\alpha + \beta) e_G^{\alpha+\beta-1}}{2A(w_G + ke_G^{\alpha+\beta} - \alpha ke_G^{\alpha+\beta})} \right) \]

Thus

\[ k(\alpha + \beta) e_G^{\alpha+\beta-1} = \lambda \iff e_G = \left( \frac{k(\alpha + \beta)}{\lambda} \right)^{1-\alpha-\beta} \quad (6) \]

In this case human capital is independent of the amount of natural resources.

**Proposition 2** If the government wins the election and revolution is no threat, human capital does not depend on natural resources.

**Case (ii): constrained solution**

Since the no-revolution constraint binds, the government maximizes:

\[
\max_{e_G, w_G} \left( \frac{W/n - w_G - \lambda e_G + \alpha ke_G^{\alpha+\beta}}{2A} \right) \times \min \left\{ \max \left\{ 0, 1 - \gamma e_G \right\} \right\}
\]

subject to

\[ w_G = -k(1 - \alpha) e_G^{\alpha+\beta} + \gamma(e_G) W/n \exp \left( -\frac{c}{q} \right) \quad (7) \]

or equivalently

\[
\max_{e_G} \left( \frac{W/n}{2A} \left( 1 - \gamma(e_G) \exp \left( -\frac{c}{q} \right) \right) + ke_G^{\alpha+\beta} - \lambda e_G \right) \times \min \left\{ \max \left\{ 0, 1 + \frac{\ln \gamma(e_G) + \ln W/n - \frac{c}{q} - U_F}{2A} \right\} \right\}
\]
The first order conditions are \( G'(e_G) \equiv -\frac{W}{n} \left( \gamma'(e_G) \exp \left( -\frac{c}{q} \right) \right) + k\left( \alpha + \beta \right)e^{\alpha + \beta - 1} - \lambda \)

\[
G'(e_G) \equiv \left[ -\frac{W}{n} \left( \gamma'(e_G) \exp \left( -\frac{c}{q} \right) \right) + k\left( \alpha + \beta \right)e^{\alpha + \beta - 1} - \lambda \right] 
\times \left( \frac{1}{2} + \frac{\ln \gamma(e_G) + \ln \frac{W}{n} - \frac{c}{q} - U_F}{2A} \right) 
+ \left( \frac{\gamma'(e_G)}{2A\gamma(e_G)} \right) \left[ 1 - \gamma(e_G) \exp \left( -\frac{c}{q} \right) \right] + ke^{\alpha + \beta} - \lambda e_G 
= 0
\]

which implies that

\[
\frac{\partial e_G}{\partial W/n} = \left( \frac{\gamma'(e_G) \exp \left( -\frac{c}{q} \right)}{G''(e_G)} \left( \frac{1}{2} + \frac{\ln \gamma(e_G) + \ln \frac{W}{n} - \frac{c}{q} - U_F}{2A} \right) \right) 
- \frac{1}{G''(e_G)} \left( \frac{1}{2A\gamma(e_G)} - \frac{\partial U_F}{\partial W/n} \frac{1}{2A} \right) 
\times \left[ -\frac{W}{n} \left( \gamma'(e_G) \exp \left( -\frac{c}{q} \right) \right) + k\left( \alpha + \beta \right)e^{\alpha + \beta - 1} - \lambda \right] 
- \left( \frac{1 - \gamma(e_G) \exp \left( -\frac{c}{q} \right)}{G''(e_G)} \right) \left( \frac{\gamma'(e_G)}{2A\gamma(e_G)} \right)
\]

where

\[
\frac{\partial U_F}{\partial W/n} = \frac{\delta(e_F)}{\delta(e_F)W/n - \lambda e_F + ke^{\alpha + \beta}}
\]

\(^{25}\)When calculating the first order conditions we implicitly assume that \( \frac{1}{2} + \frac{U_G - U_F}{A} < 1 \). If the expression becomes bigger than 1, the government wins the elections for sure. In this case it only has to take the no-revolution constraint into account.
so that

\[
\frac{\partial e_G}{\partial W/n} = \frac{\gamma'(e_G) \exp \left(-\frac{c}{q}\right)}{G''(e_G)} \left(\frac{1}{2} + \frac{\ln \gamma(e_G) + \ln W/n - \frac{c}{q} - U_F}{2A}\right) - \frac{1}{G''(e_G)} \left(\frac{1}{2A W/n} - \frac{\delta(e_F) W/n - \lambda e_F + k e_F^{\alpha+\beta}}{2A}\right)
\]

\[
\times \left[-W/n \left(\gamma'(e_G) \exp \left(-\frac{c}{q}\right) + k(\alpha + \beta) e_G^{\alpha+\beta-1} - \lambda\right) + \left[1 - \gamma(e_G) \exp \left(-\frac{c}{q}\right)\right] \left(\frac{\gamma'(e_G)}{2A\gamma(e_G)}\right)\right]
\]

(8)

**Proposition 3** If revolution is a threat human capital is likely to decrease in natural resources. It will decrease for sure if there is sufficient uncertainty in the electoral process (high $A$) or if uneducated citizens are reasonably good at managing natural resources (high $\gamma(0)$) or if the opposition is very weak ($\delta(e_F)$ low and with a low upper bound).

**Proof.** See appendix. □

Notice that the proof derives exact conditions (15 and 16) for high $A$ or high $\gamma(0)$. These conditions are sufficient but not necessary for $\frac{\partial e_G}{\partial W/n} < 0$.\(^{26}\) It is easy to see that both conditions (15) and (16) are more easily satisfied for higher revolutionary success probabilities $q$ and lower revolutionary costs $c$. Moreover, returns to human capital enter condition (15). The higher these returns (higher $\alpha + \beta$ or higher $k$) the tighter condition (15). In other words, as we will see in the next subsection everything that favors revolution loosens the conditions.

### 2.3 When does the government worry about revolution?

One thing that remains unclear in the previous exposition is the conditions under which the no-revolution constraint is binding. We now explore this issue.

Rewriting the first order conditions for the unconstrained solution allows us to calculate unconstrained $w_G$.

\[
0 = -\left(\frac{1}{2} + \frac{\ln(w_G + k(1 - \alpha)e_G^{\alpha+\beta}) - U_F}{2A}\right) + \frac{W/n - w_G - k\beta e_G^{\alpha+\beta}}{2A(w_G + k(1 - \alpha)e_G^{\alpha+\beta})}
\]

\(^{26}\)If these conditions are violated, we cannot sign $\frac{\partial e_G}{\partial W/n}$ analytically but in all simulations we have undertaken when our assumptions are violated we still observed $\frac{\partial e_G}{\partial W/n} < 0$.  

18
The solution is:

\[
 w_G = \exp \left( \text{LambertW} \left( \left( W/n + (1 - (\alpha + \beta))ke_G^{\alpha+\beta} \right) \exp^{A-U_F-1} - A + U_F + 1 \right) - k(1 - \alpha)e_G^{\alpha+\beta} \right) \tag{9}
\]

Where LambertW(.) is the Lambert W function. The direct transfer \( w_G \) we obtained assuming the no-revolution constraint is not binding satisfies that constraint when it is bigger than the constrained direct transfer given by equation (7), hence when the following inequality is true.

\[
 w_G \geq \left( \gamma(e_G)W/n \right) \exp \left( -\frac{c}{q} \right) - k(1 - \alpha)e_G^{\alpha+\beta} \equiv NR(e_G) \tag{10}
\]

To get some insight about when revolution is a concern for the government, we perform some numerical simulations, using condition (10). In those simulations we will always vary the value of \( W/n \) and some other exogenous variable simultaneously. Similarly, the figures we show depict the value of \( w_G \) and of \( NR(e_G) \), as a function of \( W/n \) and some other exogenous variable. We group these other exogenous variables into four categories depending on their economic meaning. For the simulations we use the functions\(^{28}\) \( \gamma(e) = 10^{-4} + e^2 \), and \( \delta(e) = \hat{\delta}e^{\alpha+\beta} \). The basic parameters, which are then varied individually (along with \( W/n \)) to observe the different comparative statics, are: \( (\alpha, \beta, \hat{\delta}, \lambda, A, c/q, k) = (0.5, 0.2, 0.15, 1, 1, 1, 10) \).

1. The variables \( \frac{c}{q} \) and \( \gamma(e_G) \) determine the strength of the threat of revolution. The larger is this threat, the more likely is the no-revolution constraint to bind. In other words, as the citizens become better at managing natural resources (high \( \gamma(e_G) \) for all \( e_G \)), the no-revolution constraint becomes more relevant. Similarly, for low values of \( \frac{c}{q} \) (the cost of revolution is low and/or the probability of success is high) the no-revolution constraint will always bind. When \( \frac{c}{q} \) increases, low values of \( W/n \) give rise to the unconstrained solution while the constraint binds for high values of \( W/n \). Given that revolution is already costly, it is only worthwhile if there is a lot to gain (high \( W/n \)). For sufficiently high \( \frac{c}{q} \) revolution is never an issue; it is simply too costly or too unlikely to be successful.

Figure 1 shows the impact of \( \frac{c}{q} \) on both \( w_G \) and of \( NR(e_G) \) and illustrates graphically the previous discussion.

\(^{27}\)The Lambert W function, also called the Omega function or product log, is the inverse function of \( f(w) = w \exp^w \).

\(^{28}\)We tried other functional forms, in particular \( \delta(e) = \hat{\delta} \), and the qualitative results in terms of comparative statics are similar.
2. The variables $k$, $\lambda$, $\alpha$ and $\beta$ determine the returns and costs of investment in human capital.

(a) The effect of a change in $k$, which increases (linearly) the marginal return to human capital, depends crucially on the function $\gamma(e)$. An increase in $k$, leads to higher $e$ and thus an increase in $\gamma(e)$. Both the unconstrained transfer $w_G$ and $NR(e_G)$ increase with $k$ (and with $W/n$). Whether or not the latter increases more strongly, depends on $\gamma(e)$.

i. For low $\gamma(e)$, the no-revolution constraint never binds. Citizens are simply too bad at managing natural resources.

ii. Suppose $\gamma(e)$ is sufficiently large. Then, if $k$ or $W/n$ are sufficiently low, the no-revolution constraint never binds. However, if both $k$ and $W/n$ are sufficiently high, the constraint binds. The reason for this is that the average slope of $NR(e_G)$ with respect to both $k$ and $W/n$ is higher than that of $w_G$. To understand this, notice that on the right hand side of equation (10) we have the term $\gamma(e_G)W/n$. This means that the human capital of the government $e_G$, and natural resources $W/n$ are complements in the technology for revolutions, so a simultaneous increase of $k$, and thus $e_G$, and $W/n$ are bound to have a higher effect on the possibility of revolutions than on $w_G$.

Figure 2 shows the impact of $k$ on both $w_G$ and of $NR(e_G)$ and illustrates graphically the previous discussion.

(b) The parameter $\lambda$ measures the individual’s marginal cost to acquire human capital. The effects of changing $\lambda$ are, thus, the reverse effects of changing $k$ (which, remember, is a proportionality constant on human capital returns). More precisely:

i. For low $\gamma(e)$ the no-revolution constraint never binds.

ii. If $\gamma(e)$ is sufficiently large, the no-revolution constraint binds when both $\lambda$ and $W/n$ are sufficiently low and does not bind if either $\lambda$ or $W/n$ are sufficiently high.

Figure 3 shows the impact of $\lambda$ on both $w_G$ and of $NR(e_G)$.

(c) The parameters $\alpha$ and $\beta$ determine the returns to scale of human capital. We assume $\alpha + \beta < 1$, hence returns to scale will always be decreasing. Since both parameters have the same qualitative effect, we will describe only the effects of $\alpha$. For low $\alpha$, the no-revolution constraint always binds except for very high values of $W/n$. When $\alpha$ increases, the fraction of values of $W/n$ in which the unconstrained solution holds
increases, until $\alpha$ is so high that only the unconstrained solution holds. The intuition is as follows: returns to human capital are not relevant when revolution is successful since there is no productive activity of workers in case of revolution. Hence, the revolution is more attractive for low values of $\alpha$. When $\alpha$ increases, it is more costly to forgo the returns from productive activity, and revolution will only be attractive if there are sufficient natural resources to be managed. For high enough $\alpha$, it is simply too costly not to engage in the productive activity, hence revolution is never an issue.

Figure 4 shows the impact of $\alpha$ on both $w_G$ and of $NR(e_G)$ and illustrates graphically the previous discussion.

3. The function $\delta(e)$ is a measure of institutional quality. Better political institutions (higher $\delta(e)$ functions) allow the fringe to offer a higher utility $U_F$ to voters. Hence, the government has to react with a higher direct transfer $w_G$ which implies that the no-revolution constraint will bind less often. In other words, with good institutions, revolution will not occur.

4. The aggregate shock $A$ to voters’ preferences measures the extent to which policies matter for winning the elections. The bigger the shock, the less important are the promised utilities to voter. For very low $A$, we always have the unconstrained solution. When $A$ increases, the constraint soon bites and we only get the unconstrained solution for low $W/n$. The higher $A$, the smaller the fraction of value for which the unconstrained solution holds. This happens because $w_G$ decreases with $A$, since promised utilities have a smaller effect on the probability of winning the elections, while $NR(e_G)$ is independent of $A$.

Figure 5 shows the impact of $A$ on both $w_G$ and of $NR(e_G)$ and illustrates graphically the previous discussion.

### 2.4 Determining the winner of the elections

The probability that the government wins the election is directly related to $U_G - U_F$. To gain some insight we will discuss the case when the no-revolution constraint does not bind. From (9) we can conclude that

$$U_G - U_F = \text{LambertW}
\left((W/n + (1 - (\alpha + \beta))ke_G^{\alpha+\beta}) \exp^{A-U_F-1}\right) - A + 1$$

Since the LambertW function is increasing we only have to look at the derivative of its argument. Thus we have
$$\frac{\text{sign}}{\partial} \left( \frac{U_G - U_F}{W/n} \right) = \text{sign} \left( \exp^{A-U_F-1} \left( 1 - \left( W/n + (1 - (\alpha + \beta))k e_G^{\alpha+\beta} \right) \frac{\partial U_F}{\partial W/n} \right) \right)$$

$$= \text{sign} \left( 1 - \frac{(W/n + (1 - (\alpha + \beta))k e_G^{\alpha+\beta})\delta(e_F)}{\delta(e_F)W/n - \lambda e_F + k e_F^{\alpha+\beta}} \right)$$

(11)

Whether this sign is positive or negative, hence whether the probability that the government wins the election is increasing or decreasing is generally going to depend on the parameters of the model. However, a couple of things can be deduced from this expression. For \( W/n = 0 \) we know from equation (5) that

\[ e_F = \left( \frac{(\alpha+\beta)k}{\lambda} \right)^{\frac{1}{1-\alpha-\beta}} \]

Clearly, if \( \frac{(\alpha+\beta)k}{\lambda} \) is low enough, (11) is positive. On the other hand, for very large \( W/n \) when the variation of \( e_F \) is smaller than that of \( W/n \) then (11) will asymptote to zero. From this argument it is not clear whether it could ever be decreasing. To confirm that in fact it can, we perform a numerical simulation using the same basic parameter values and functional forms as in subsection 2.3. The result of this simulation is shown in Figure 6. The figure displays the two features we uncovered analytically and also shows that for sufficiently high \( W/n \) the sign is negative.

The fact that the derivative can be both positive and negative reflects that two economic forces are at work. On the one hand, as resources increase, the government can pay higher direct transfers \( w_G \), thus increasing its chances of winning. On the other hand, the fringe can also offer better terms, especially through the channel of human capital \( e_F \), which also enhances its probability of winning and makes the fringe a better administrator of natural resources. The concavity of the effort function makes it more likely that the first effect dominates in the beginning. The effect of direct transfers hits the margin directly from the beginning, whereas the effect of human capital needs more natural resources to have the same marginal impact.

Clearly, by bounding the \( \delta(e) \) function one could ensure that (11) is never negative, which seems to be the relevant case according to the empirical evidence. Recall that we interpreted the \( \delta(e) \) function as a proxy for institutional quality. \( \delta(e) \) low and bounded corresponds to a country where institutions are weak and the Fringe cannot manage natural resources as efficiently as the government even for high levels of education. In this case, any natural resource finding increases the chances of the incumbent government to stay in power. Only strong institutions make it less likely that natural resources will allow the incumbent to become more entrenched.
3 Empirical evidence

In order to test our theory empirically we first derive an econometric model for our main hypothesis. Our claim is as follows: the level of institutional quality $d_i$ of a country $i$ affects how human capital $H_i$ changes with natural resources $r_i$. For low levels of institutional quality human capital decreases with natural resources, while for high level of institutional quality human capital increases with natural resources. Mathematically, our main hypothesis can be formulated as

$$\frac{\partial H_i}{\partial r_i} = f_{Z_i}(d_i) \text{ with } f_{Z_i}(0) < 0 \text{ and } f_{Z_i}(d) > 0 \text{ for } d \text{ high enough} \quad (12)$$

and $f'_{Z_i} > 0$ where $Z_i = \begin{bmatrix} z_i^1 \\ \vdots \\ z_i^n \\ d_i \end{bmatrix} = \begin{bmatrix} Z_i^{-d} \\ d_i \end{bmatrix}$ and $Z_i^{-d}$ captures all other potentially relevant variables. To derive an econometric model we will assume that

$$f_Z(d) = \alpha_0 + \alpha_1 d$$

i.e. $f_Z(\cdot)$ is a linear function of $d_i$ and it is independent of $Z$. With this assumption our main hypothesis (12) can be stated as

$$\alpha_0 < 0 \text{ and } \alpha_1 > 0. \quad (13)$$

If we further assume that $H_i$ depends linearly on the $Z_i$’s our econometric model can be formulated as

$$H_i = \delta'Z_i + (\alpha_0 + \alpha_1 d_i) r_i + \varepsilon_i \quad (14)$$

3.1 Data

We use several sources of data. For natural resources we take natural capital share ($\text{naturalk}$) from Gylfason and Zoega (2006) who constructed this measure from World Bank Data. Natural capital is the sum of “subsoil wealth”, timber, non-timber benefits of forests, cropland, pasture land, and the opportunity cost of protected areas. In turn, subsoil wealth is the present value of a constant stream of economic profits on “resource rents” on various fuels and minerals; that is, gross profit on extraction less depreciation of capital and normal return on capital. As a proxy for the quality of political institutions ($\text{polriginv}$) we use the inverse of the Gastil Index of Political
Rights constructed by the Freedom House. Thus, in our regressions political rights are measured on a one-to-seven scale, with one representing the lowest degree of freedom and seven the highest. Political Rights’ data for a specific year is the previous five years’ average. Human capital will be measured by primary school enrollment. The source for the data on primary school enrollment is the World Development Indicators from UNESCO. The additional controls are taken from the World Development Indicators (WDI) and are the log of GDP (lgdp), measures of fertility, mortality and birth rates and the pupil teacher ratio in primary schools (pupil). Natural capital is available for 1994. All other variables are typically available every 5 years from 1970 to 1995. In the regression we use per country averages for this period and only include countries which have at least 3 observations for all variables. This makes 78 countries. Therefore

\[
Z = \begin{bmatrix}
    \text{lgdp} \\
    \text{fertility} \\
    \text{mortality} \\
    \text{pupil} \\
    \text{birth} \\
    \text{constant} \\
    \text{polrigniv}
\end{bmatrix}
\]

when we use all variables, \( d \) is polrigniv and \( r \) is naturalk.

### 3.2 Results

We now report various estimates of the parameters \( (\delta', \alpha_0, \alpha_1) \). In all tables the coefficients estimated correspond to each variable, with cross2 representing \( \alpha_1 \) which multiplies \( d_ir_i \). We first estimate (14) by OLS. The results are reported in table 1. Most importantly, natural capital has

---

29 For a detailed explanation of the index and the methodology see http://www.freedomhouse.org/template.cfm?page=35&year=2005
30 The countries are: Austria, Bangladesh, Belgium, Benin, Botswana, Brazil, Burkina Faso, Burundi, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Congo, Rep., Costa Rica, Cote d’Ivoire, Denmark, Dominican Republic, Ecuador, Egypt, Arab Rep., El Salvador, Finland, France, Gambia, The Ghana, Greece, Guatemala, Haiti, Honduras, India, Indonesia, Ireland, Italy, Jamaica, Japan, Jordan, Kenya, Korea, Rep., Madagascar, Malawi, Malaysia, Mali, Mauritania, Mauritius, Mexico, Morocco, Mozambique, Nepal, Netherlands, Nicaragua, Niger, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Portugal, Rwanda, Senegal, Sierra Leone, South Africa, Spain, Sri Lanka, Sweden, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, United Kingdom, United States, Uruguay, Venezuela, Zambia and Zimbabwe.
a negative and significant effect and the cross variable has a positive and significant effect. T-tests show that hypothesis $\alpha_0 = 0$ and $\alpha_1 = 0$ are each rejected at 99% confidence level. This means that the higher the institutional quality, the lower the negative impact of natural resources. In fact, when the index for political rights is above $-\alpha_1/\alpha_2 = 3.397$ the net effect is positive. This is consistent with the predictions of our model. Other controls also yield sensible estimates: log GDP has a positive effect on human capital, fertility and mortality have negative effects.

Table 1: OLS regression.

|        | Coeff  | Std. Err. | t     | P(\(|\beta_i| > 0\)) |
|--------|--------|-----------|-------|-----------------------|
| naturalk | -1.153341 | .4238907 | -2.72 | 0.008                 |
| cross2  | .3395294  | .1244045 | 2.73  | 0.008                 |
| lgdp    | 3.541507   | 2.756116 | 1.28  | 0.203                 |
| fertility | -20.29832 | 5.814107 | -3.49 | 0.001                 |
| mortality | -.2884667 | .0950492 | -3.03 | 0.003                 |
| pupil   | .2558133 | .1964023 | 1.30  | 0.197                 |
| birth   | 2.493832   | .8983785 | 2.78  | 0.007                 |
| polriginv | -7.585228 | 2.195964 | -3.45 | 0.001                 |
| constant | 116.8061   | 24.83554 | 4.70  | 0.000                 |

However, OLS regressions are vulnerable to an endogeneity problem: institutional quality might be endogenous to human capital and therefore correlated with the shock $\varepsilon$. If this is so, OLS is inconsistent. To get a consistent estimator we need an instrument for institutional quality which is orthogonal to the residual of the main equation (14). We propose to use the average of the levels of institutional quality of the neighboring countries as an instrument.\(^{31}\) We define neighbors as follows: (i) if country A is not an island country, country A is a neighbor of country B if they have a common border. (ii) if country A is an island country, neighbors are defined in terms of physical closeness. For small island countries we only consider the closest neighbor.\(^{32}\)For big

---

\(^{31}\)Another possibility would be to use lagged levels of institutional quality. This method is used by Egorov et al. (2006) who use the Polity IV dataset as a proxy for democracy and replace their contemporaneous measure of democracy (averaged 1993-2003) with its average level in 1980-1992. We prefer our method since it allows us to use more years of observations.

\(^{32}\)The small island countries used in the regressions are: Mauritius, Trinidad and Tobago.
island countries the four nearest neighboring countries are considered. The threshold surface was 10,000 square kilometers. The correlation between the instrument and a country’s own political right index “polriginv” is 0.6637.

Notice that institutional quality enters the regression twice, since it enters by itself and also as the cross variable $d_i r_i$. Hence, we need an instrument for the cross variable, too. To introduce as little noise as possible we propose the following two-stage-Non-linear-least square (2SNLLS) estimator. We first regress polriginv on the exogenous variables $Z_i^{-d}$ and the average institutional quality value of the neighbors. Then we construct the calculated values of institutional quality and replace $d_i$ by these calculated values in the right hand side of (14) both in linear and cross-product elements. The asymptotic variance-covariance matrix of the estimator cannot be done by standard two-stage least squares because the presence of the cross product $d_i r_i$ introduces heteroskedasticity in the second stage regression. Since standard econometric texts are of no help, we prove in the appendix that the proposed estimator is consistent and we find the asymptotic distribution by GMM. This allows us to build the corresponding t-tests. The regression results we report below are based on the t-tests derived in the Appendix. Our first 2SNLLS regression (table 2) uses the same exogenous variables as the OLS regression (with institutional quality replaced by its instrument).

In this 2SNLLS all estimates have the same sign as in the OLS regression but their significance is reduced. It is standard in instrumental variable that variance is higher. We now follow the standard procedure and drop insignificant variables from the regression, namely GDP, polriginv and pupil.

By construction there is multicollinearity between polriginv and cross2, hence it is natural that we drop polriginv, thereby avoiding multicollinearity. This new regression is reported in table 3. All variables are now significant and our main hypothesis is confirmed. Natural capital has a negative and significant effect (at a 1% level) and the cross variable has has a positive and significant effect (at a 1% level). When the index of political rights is above $-\alpha_1/\alpha_2 = 3.657$ the net effect is positive, consistently with our model.

To summarize, the above regressions clearly corroborate our theory: the quality of institutions

33The big island countries used in the regressions are: Dominican Republic, Jamaica, Japan, Madagascar, Philippines, Sri Lanka

34Notice, though, that our main prediction is confirmed already in this regression as cross2 is significant and it has the correct size and sign.

35Of the countries used in our regression 47 countries have an index above 3.657.
Table 2: 2SNLLS regression.

|        | Coefficient | Std. Error | t-value | P(|β_i| > 0) |
|--------|-------------|------------|---------|-------------|
| naturalk | -1.3655435  | .7105      | -1.922  | 0.0273      |
| cross2  | .38152022   | .19827     | 1.9242  | 0.0272      |
| lgdp    | .50337745   | 6.4878     | 0.0776  | 0.4690      |
| fertility| -16.590448  | 7.6319     | -2.1738 | 0.0149      |
| mortality| -2.9344412  | 0.1005     | -2.92   | 0.0018      |
| pupil   | .30395595   | 0.25787    | 1.1787  | 0.1193      |
| birth   | 1.9962673   | 1.02362    | 1.9502  | 0.0256      |
| polriginv| -4.5996464  | 8.8332     | -0.5207 | 0.3013      |
| const   | 125.3595    | 28.74004   | 4.362   | 0.0000      |

is decisive in determining whether natural resources are a blessing or a curse. In countries with good institutions natural resources enhance education. In countries with bad institutions natural resources are detrimental to education.\footnote{If this dependence (i.e. the crossproduct) would be ignored, we would wrongly conclude from the evidence that natural resources can only be a curse. We looked at different regressions without the cross product. In all of them the coefficient for natural capital was negative, but it was not always significant.} In all regressions the cutoff value for our index of political rights is close to 3.5. Hence, in all countries clearly classified as free by the Freedom House (which correspond to our index 5.5 to 7) natural resources are a blessing and are a curse in all countries the Freedom House classifies as not free (which correspond to our index 1 to 2.5). The turning point lies in the partly free countries (with an index 3 to 5).

4 Conclusion

In this paper we have presented a formal political-economy analysis of the impact of natural resources on human capital accumulation. In our model, citizens exert control over politicians via an election and can always initiate a revolution if they are dissatisfied with the proposed policies. Since it is a well-documented fact that natural resources have led to civil unrest, it is important to incorporate this possibility into the model. To our knowledge this is the first paper to allow
simultaneously for political competition, elections and revolution.\textsuperscript{37,38} We propose to model the possibility of revolution by introducing a new constraint into the model, which we denote the \textit{no-revolution constraint}. Under this constraint politicians select their policies so that there are no sufficiently large sectors of the population who want to block this policy by starting a revolution. In the context of natural resources, this constraint can be taken literally. However, we would like to emphasize that this constraint might be introduced in many other models: the economic literature is full of policy recommendations which no sane politician has dared to implement even if a majority of the population would benefit from them. This sounds contrary to both economic and political theory, but we would argue that there are good practical reasons for the outcome that the models overlook.

These policy recommendations arise in models where the policy resulting from the voting mechanism (e.g. the policy preferred by the median voter) would harm a sizable proportion of

\textsuperscript{37}Introducing revolution in political economy models is not an innovation “per se.” Acemoglu and Robinson (2001) explain the “extension of the franchise” in precisely this way. But notice that in their work, revolution is a threat from citizens “excluded” from the vote, who thus have no alternative. In our work, “revolution” is an added tool for all citizens, not an alternative when there is not a chance to vote. Acemoglu and Robinson (2006) do include the possibility of revolting in democracy. However, this possibility does not operate as a constraint for the government. It is simply a binary choice for the poor (already the median voter and thus the tax setters in the democracy).

\textsuperscript{38}There is a growing theoretical literature relating power struggles to natural resources, however in this literature people either have no democratic control over rulers (e.g. Olsson 2007 who sets up a predator-prey model in which rebels choose between peaceful production and predation on natural resources controlled by the ruler, or Wick 2008 using a Stackelberg model with limited endowments), or it is the political elite which is initiating the revolution (e.g. Aslaksen and Torvik 2006, Caselli 2006).
the population. Such policies are not implemented because the sector that would be harmed has pressure instruments on top of their votes to block them, and these pressure instruments can be modeled by the no-revolution constraint.\footnote{One example of such a policy recommendation is the abolition of capital taxes. Lucas (1990) has shown that the optimal capital tax is zero. It has also been shown that the representative consumer would vote for a capital tax of zero. Even in a model with heterogeneous agents (Garcia-Milà, Marcet and Ventura, 2001) the median voter is likely to vote in favor of abolishing capital taxes. This, however, can harm as much as a third of the population. This part of the population would probably go to great lengths in order to avoid the zero capital tax.} Hence, the importance of our proposed modeling innovation lies far beyond the topic studied.

In terms of the topic we study, our contribution is to incorporate simultaneously the five empirical facts on natural resources presented in the introduction.\footnote{In recent years the consensus on these facts has started to fade. E.g. Brunnschweiler and Bulte (2008) argue that while there is correlation between resources and conflict on the one hand and slow growth on the other, the causality is reverse than traditionally understood: conflict and bad institutions make countries depend on primary exports and not the other way round. Which conclusion is reached depends on the measure of resource dependence used. In our regressions we use the natural capital share from Gylfason and Zoega (2004) which typically gives rise to the resource curse.} We can explain when natural resources are a blessing and when they are a curse (Fact 2) and we capture the importance of the quality of institutions (Fact 3). A further result of our model is that natural resources may be bad for political turnover and will be so in countries with bad political institutions (Fact 4). In those countries natural resources strengthen the position of the incumbent government, who typically chooses policies which do not enhance, or are even detrimental, to human capital accumulation and therefore growth. If the majority of countries with natural resources have bad institutions we can expect that natural resources are bad for growth on average (Fact 1: \textit{the curse}). Our model also answers the question of when natural resources lead to a revolution: in countries with bad political institutions (Fact 5).\footnote{Observe that in our model revolution never occurs in equilibrium. Nevertheless our model can guide us about the possibility of revolution: obviously adding some noise would lead to occasional violations of the no revolution constraint and result in a revolution. According to our model this can only happen in countries with bad institutions.}

Our model links natural resources to education which is an established engine for growth. Nevertheless, we want to be sure that the empirical facts which are stated in terms of growth are also valid if we use education. We therefore tested these facts for education in cross-country regressions, and we find that they indeed hold.

Some authors have suggested that the size of a country matters for the effect of natural resources. This is captured in our model, where country size is measured by $n$. Increasing $n$ has the
same effect as decreasing natural resources $W$.

In our model, the income of the government stems only from natural resources. In a more complete model the government can also receive income by taxing productive activity. This is one of the extensions we would like to study in the future. The existence of productive activity has an effect on the incentives of politicians to encourage human capital accumulation: better education should enhance productive activity, which in turn enables the government to extract more taxes. But better education also strengthens the opposition and the ability of citizens to engage in a successful revolution. We expect that the incumbent government will prefer not to enhance education, since education weakens its political position and it is easier for them to increase their income from natural resources than by taxing productive activity. Natural resources are easily appropriated by corrupt politicians. So are some unnatural resources, like foreign aid. Is there a link between natural resources and foreign aid? Can our model make predictions about the effects foreign aid might have on education or growth?

The answer is yes. Once the foreign aid is granted it is very difficult for international institutions to avoid that politicians steal foreign aid. Empirical evidence suggests that only a small percentage of the aid actually reaches its desired objective. In Uganda only 13% of foreign aid granted for education in 1991-1995 actually reached primary schools (Reinikka and Svensson (2004)). The evidence for other African countries is similar. As with natural resources the quality of institutions is crucial in limiting stealing from foreign aid. But similarly to natural resources, foreign aid tends to be detrimental to democracy: studying 108 recipient countries of foreign aid in the period 1960 to 1999 Djankov, Montalvo and Reynal-Querol (2005) find a negative effect of foreign aid on democracy which is much bigger than the negative effect of natural resources. Like natural resources foreign aid can be the cause of civil war and revolution.\footnote{Maren (1997) provides evidence that the cause of the civil war in Somalia was the control over foreign aid.}

Given these empirical similarities between the effects of natural resources and foreign aid, we can use our model to make predictions about when foreign aid is a blessing and when it is a curse. In countries with good institutions, foreign aid will enhance growth, while the opposite will happen in countries with bad institutions. Typically it is the latter group of countries that receives foreign aid. Our model recommends that only poor countries that have good institutions should be granted foreign aid.
A Proofs

Proposition 3 If revolution is a threat human capital is likely to decrease in natural resources. It will decrease for sure if the there is sufficient uncertainty in the electoral process (high $A$) or if uneducated citizens are reasonably good at managing natural resources (high $\gamma(0)$) or if the opposition is very weak ($\delta(e_F)$ low and with a low upper bound).

Proof. Equation (8) gives us the expression for $\frac{\partial G}{\partial W/n}$. We know that: $G''(e_G) \leq 0$ to guarantee the satisfaction of second order conditions. Also, $\left(\frac{1}{2} + \frac{\ln \gamma(e_G) + \ln W/n - \frac{c}{q} - U_F}{2A}\right) \geq 0$, since it is a probability. Hence the first line of the expression is negative.

Claim 1 $[-W/n \left(\gamma'(e_G) \exp \left(-\frac{c}{q}\right)\right) + k(\alpha + \beta)e_G^{\alpha+\beta-1} - \lambda] < 0$.

Proof.

$$\left[-W/n \left(\gamma'(e_G) \exp \left(-\frac{c}{q}\right)\right) + k(\alpha + \beta)e_G^{\alpha+\beta-1} - \lambda\right] + \left(\frac{\gamma'(e_G)}{2A\gamma(e_G)}\right) \left[W/n \left(1 - \gamma(e_G) \exp \left(-\frac{c}{q}\right)\right) + ke_G^{\alpha+\beta} - \lambda e_G\right] = 0$$

but

$$\left(\frac{\gamma'(e_G)}{2A\gamma(e_G)}\right) \left[W/n \left(1 - \gamma(e_G) \exp \left(-\frac{c}{q}\right)\right) + ke_G^{\alpha+\beta} - \lambda e_G\right] > 0$$

since $W/n \left(1 - \gamma(e_G) \exp \left(-\frac{c}{q}\right)\right) + ke_G^{\alpha+\beta} - \lambda e_G$ are profits of incumbent and $\gamma'(e_G) > 0$, and the result follows.

Claim 2

$$\frac{1}{2AW/n} - \frac{\delta(e_F)}{\delta(e_F)W/n - \lambda e_F + ke_F^{\alpha+\beta} \frac{1}{2A}} > 0$$

Proof.

$$\frac{1}{2AW/n} - \frac{\delta(e_F)}{\delta(e_F)W/n - \lambda e_F + ke_F^{\alpha+\beta} \frac{1}{2A}} = \frac{1}{2A} \left(\frac{1}{W/n} - \frac{1}{\delta(e_F)} \left(k e_F^{\alpha+\beta} - \lambda e_F\right)\right)$$

and $(ke_F^{\alpha+\beta} - \lambda e_F) > 0$ if individual investment in human capital is possible.

Claims (1) and (2) show that the second/third line of (8) is also negative. Unfortunately the fourth line is unambiguously positive since $\left(1 - \gamma(e_G) \exp \left(-\frac{c}{q}\right)\right) > 0$ and $\gamma'(e_G) > 0$. 31
To try to sign this fourth line, let’s first bound the first line of (8). If \( \delta(e_F) \) were equal to 1, i.e. the fringe is as efficient as the government in managing natural resources, then from (5)

\[
\left( \frac{(\alpha + \beta)k}{\lambda} \right)^{\frac{1}{1-(\alpha+\beta)}} = e_F^x
\]

Thus

\[
U_F \leq \ln \left( \frac{W/n - \lambda \left( \frac{(\alpha + \beta)k}{\lambda} \right)^{\frac{1}{1-(\alpha+\beta)}} + k \left( \frac{(\alpha + \beta)k}{\lambda} \right)^{\frac{\alpha+\beta}{1-(\alpha+\beta)}}}{1} \right)
\]

Thus

\[
\frac{\ln \gamma(e_G) + \ln W/n - \frac{\xi}{q} - U_F}{2A} \geq \frac{1}{2A} \ln \left( \frac{\gamma(e_G) \exp \left( -\frac{\xi}{q} \right) \frac{W/n}{1} - \lambda \left( \frac{(\alpha + \beta)k}{\lambda} \right)^{\frac{1}{1-(\alpha+\beta)}} + k \left( \frac{(\alpha + \beta)k}{\lambda} \right)^{\frac{\alpha+\beta}{1-(\alpha+\beta)}}}{1 - \lambda \left( \frac{(\alpha + \beta)k}{\lambda} \right)^{\frac{1}{1-(\alpha+\beta)}} + k \left( \frac{(\alpha + \beta)k}{\lambda} \right)^{\frac{\alpha+\beta}{1-(\alpha+\beta)}}} \right)
\]

Assume

\[
\frac{1}{2A} \ln \left( \frac{\gamma(0) \exp \left( -\frac{\xi}{q} \right)}{1 - \lambda \left( \frac{(\alpha + \beta)k}{\lambda} \right)^{\frac{1}{1-(\alpha+\beta)}} + k \left( \frac{(\alpha + \beta)k}{\lambda} \right)^{\frac{\alpha+\beta}{1-(\alpha+\beta)}}} \right) \geq -\frac{1}{4} \tag{15}
\]

Then if we take the first and third terms of (8) we have

\[
\frac{\partial e_G}{\partial W/n} \leq \left( \frac{\gamma'(e_G) \exp \left( -\frac{\xi}{q} \right)}{G''(e_G)} \right)^{1/4} - \frac{1}{G''(e_G)} \left( \frac{\gamma'(e_G)}{2A\gamma(e_G)} \right) = \frac{\gamma'(e_G)}{4G''(e_G)} \left( \exp \left( -\frac{c}{q} \right) - \frac{2}{A\gamma(e_G)} \right)
\]

Assume that

\[
A\gamma(0) \exp \left( -\frac{c}{q} \right) > 2 \tag{16}
\]

and we are done. Both assumptions (15) and (16) are satisfied for \( A \) sufficiently high or \( \gamma(0) \) sufficiently high.

To study the case of a weak opposition, we have to distinguish two possible scenarios: (i) the opposition has some chance to win the elections and (ii) the government wins the election for sure.
In the first case notice that we could use the upper bound on $\delta(e_F)$ to bound the first line of (8). Let this bound be $\delta$ and use it to calculate an upper bound for $U_F$ as before. This gives us

$$U_F \leq \ln \left( \delta W/n - \lambda \left( \frac{\alpha + \beta}{\lambda} \right)^{\frac{1}{1-(\alpha+\beta)}} + k \left( \frac{\alpha + \beta}{\lambda} \right)^{\frac{\alpha + \beta}{1-(\alpha+\beta)}} \right)$$

Thus

$$\frac{\ln \gamma(e_G) + \ln W/n - \frac{c}{q} - U_F}{2A} \geq \frac{1}{2A} \ln \left( \frac{\gamma(e_G) \exp \left( -\frac{c}{q} \right) W/n}{\delta W/n - \lambda \left( \frac{\alpha + \beta}{\lambda} \right)^{\frac{1}{1-(\alpha+\beta)}} + k \left( \frac{\alpha + \beta}{\lambda} \right)^{\frac{\alpha + \beta}{1-(\alpha+\beta)}}} \right)$$

$$= \frac{1}{2A} \ln \left( \frac{\gamma(e_G) \exp \left( -\frac{c}{q} \right)}{\delta - \lambda \frac{W/n}{\left( \frac{\alpha + \beta}{\lambda} \right)^{\frac{1}{1-(\alpha+\beta)}}} + k \frac{W/n}{\left( \frac{\alpha + \beta}{\lambda} \right)^{\frac{\alpha + \beta}{1-(\alpha+\beta)}}}} \right)$$

and we could use the following assumption (17) instead of (15).

$$\frac{1}{2A} \ln \left( \frac{\gamma(0) \exp \left( -\frac{c}{q} \right)}{\delta - \lambda \frac{W/n}{\left( \frac{\alpha + \beta}{\lambda} \right)^{\frac{1}{1-(\alpha+\beta)}}} + k \frac{W/n}{\left( \frac{\alpha + \beta}{\lambda} \right)^{\frac{\alpha + \beta}{1-(\alpha+\beta)}}}} \right) \geq -\frac{1}{4} \quad (17)$$

The lower $\delta$, the lower the necessary $A$ to satisfy assumption (17).\(^{43}\)

Let us know study the case where $U_F$ is so small that the government wins the election for sure. In this case G’s problem can be simplified to:

$$\max_{e,w} \ W/n - w - \lambda e + \alpha ke^{\alpha+\beta}$$

subject to $U(w + ke^{\alpha+\beta} - \alpha ke^{\alpha+\beta}) \geq U(\gamma(e)W/n) - \frac{c}{q}$

Assume as before that $U(x) = \ln(x)$. Then

$$\max_{e,w} \ W/n - w - \lambda e + \alpha ke^{\alpha+\beta}$$

subject to $w + ke^{\alpha+\beta} - \alpha ke^{\alpha+\beta} \geq (\gamma(e)W/n) \exp \left( -\frac{c}{q} \right)$

\(^{43}\)Notice that assumption (17) holds for any upper bound $\delta$ of the delta-function, also for values bigger than 1. In these cases higher $A$ are needed to satisfy assumption (17). However, this is not worrisome since subsection 2.2.1 shows that for high delta-function the no revolution constraint does not bind.
Thus an equivalent way of writing the problem is:

\[
\max_{e,w} \frac{W}{n} \left( 1 - \gamma(e) \exp \left( -\frac{c}{q} \right) \right) + ke^{\alpha+\beta} - \lambda e
\]

The FOC in this case are:

\[
G'(e) \equiv -\frac{W}{n} \left( \gamma'(e) \exp \left( -\frac{c}{q} \right) \right) + k(\alpha + \beta)e^{\alpha+\beta-1} - \lambda = 0
\]

\[
\frac{\partial e}{\partial W/n} = \frac{\gamma'(e) \exp \left( -\frac{c}{q} \right)}{G''(e)}
\]

\[
G''(e) = -\frac{W}{n} \left( \gamma''(e) \exp \left( -\frac{c}{q} \right) \right) + k(\alpha + \beta - 1)(\alpha + \beta)e^{\alpha+\beta-2}
\]

As before, \(G''(e) \leq 0\) in the optimum, which implies that \(\frac{\partial e}{\partial W/n} \leq 0\). This assumes that we have an interior solution. A sufficient condition for this is, \(\gamma''(\tilde{e}) \geq 0\) for \(\tilde{e}\) with \(G'(\tilde{e}) = 0\).

**B The estimator (theory)**

Here we discuss our estimation of equation (14) when we allow for the variable \(d_i\) to be endogenous and, therefore, potentially correlated with the residual \(\varepsilon_i\). From the point of view of the hypothesis to be tested, the coefficients of interest are the \(\alpha'\)s, that multiply the terms involving natural resources \(r_i\) in equation (14). But from the econometric point of view the difficulty lies with the coefficients multiplying the (potentially endogenous) variable \(d_i\), therefore we rewrite equation (14) as

\[
H_i = \tilde{\delta}'x_i + \tilde{\gamma}'(d_i, d_i r_i)' + \varepsilon_i
\]

(18)

where

\[
\delta' \equiv [\beta^{-d_i'}, \alpha_0], \quad \gamma' = [\gamma_1, \gamma_2] \equiv [\beta, \alpha_1],
\]

\[
x_i' \equiv [Z_i^{-d_i'}, r_i]
\]

so that \(x_i\) collects the \(m\) exogenous variables in the regression. \(\tilde{\delta} \in R^m\) and the scalars \(\tilde{\gamma}_i\) are true values of the parameters, we use the symbols without tilde to denote generic values for these parameters.

We have to instrument the vector \((d_i, d_i r_i)\). For this purpose we assume that \(\varepsilon\) is orthogonal to the \(x\)’s and to the neighbor’s level of natural resources in country \(i, n_i\):

\[
E(\varepsilon_i(x_i', n_i)) = 0.
\]

(19)
Note the variable \( n \) is not in the main regression equation (18) but it is assumed to be orthogonal to the residual in this equation. Our approach will allow us to instrument the two terms \((d_i, d_ir_i)\) with just one instrument.

We have \( I \) observations of all the variables. We assume that the joint distribution of all variables is identical across countries. Let us denote:

\[
X = \begin{bmatrix} x'_1 \\ \vdots \\ x'_I \end{bmatrix}, \quad Y = \begin{bmatrix} H_1 \\ \vdots \\ H_I \end{bmatrix}, \quad D = \begin{bmatrix} d_1 \\ \vdots \\ d_I \end{bmatrix},
\]

\[
N = \begin{bmatrix} n_1 \\ \vdots \\ n_I \end{bmatrix}, \quad DR = \begin{bmatrix} d_1 r_1 \\ \vdots \\ d_I r_I \end{bmatrix},
\]

The OLS estimator

\[
\begin{bmatrix} \beta^{OLS} \\ \gamma^{OLS} \end{bmatrix} \equiv [(X, D, DR)'(X, D, DR)]^{-1} (X, D, DR)'Y
\]

is potentially inconsistent because we think that \( E(\varepsilon_i(d_ir_i, d_i)) \neq 0 \).

To get a consistent estimator we propose a two-stage least squares estimator. Since the endogenous variable \( d \) enters non-linearly in (18) we call this estimator “two-stage-Non-linear-least-squares” (2SNLLS). We will see how some difficulties arise in finding the asymptotic distribution due to the fact that the cross-product term \( d_ir_i \) intrinsically introduces heteroskedasticity in the second stage regression. The gain with our approach is that we have a more efficient estimator than a version of standard 2SLS, we discuss this in the second half of this appendix.

Let us define our estimator. In a first stage we regress \( d \) on the exogenous variables

\[
d_i = (x'_i, n_i)\tilde{\phi} + u_i
\]

where \( \tilde{\phi} \) are the true projection coefficients satisfying

\[
E[(x'_i, n_i) u_i] = 0
\]

Such an \( \tilde{\phi} \) always exists as long as all variances are finite and it is consistently estimated by the OLS estimator:

\[
\phi^{OLS} = [(X, N)'(X, N)]^{-1} (X, N)'D
\]
With this we construct the calculated values of institutional quality, "replace" $d_i$ in the right hand side of (18) by these calculated values, and run a second stage regression including the linear and cross-product elements. More precisely, letting

$$
\hat{d}_i = (x'_i, n_i) \phi^{OLS}
$$

$$
\hat{D} = (X, N) \phi^{OLS}
$$

$$
\hat{D} \hat{R} = \begin{bmatrix}
\hat{d}_1 r_1 \\
\vdots \\
\hat{d}_I r_I 
\end{bmatrix}
$$

the second stage regression is

$$
\begin{bmatrix}
\hat{\gamma}_{2SNLLS} \\
\hat{\gamma}_{2SNLLS}
\end{bmatrix} = \left[ (X, \hat{D}, \hat{D} \hat{R})'(X, \hat{D}, \hat{D} \hat{R}) \right]^{-1} (X, \hat{D}, \hat{D} \hat{R})'Y
$$

In this "second stage regression" heteroskedasticity appear intrinsically. To see this, notice that we can rewrite (18) as

$$
H_i = \tilde{\gamma}_1 x_i + \tilde{\gamma}_2 r_i (x'_i, n_i) \phi + u_i + \varepsilon_i
$$

$$
= \tilde{\gamma}_1 x_i + \tilde{\gamma}_2 r_i \tilde{d}_i + \varepsilon_i + U_i
$$

The second equality gives the second stage regression, notice that the "actual" residual in this second stage regression is $U_i \equiv \varepsilon_i + u_i (\tilde{\gamma}_1 + r_i \tilde{\gamma}_2)$ which can not possibly be independent of $r_i$, hence homoskedasticity is out of the question. Larger $r$’s will imply larger variance of the residual. This means that the standard variance covariance matrix of 2SLS is not correct and we need to construct test-statistics with a more general framework.

To prove that 2SNLLS is a consistent estimator and to find the asymptotic distribution we cast this estimator in the framework of GMM. For easy reference, we describe the basic setup in GMM: given a function $g(z_i, b)$ of observables $z$ and parameters $b$, assume we know that an orthogonality condition such that

$$
E(g(z_i, \tilde{b})) = 0
$$

holds at the true value $\tilde{b}$, but

$$
E(g(z_i, b)) \neq 0 \quad \text{for } b \neq \tilde{b}
$$
Consider only the case where \( g(z_i, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n \) (equal number of orthogonality conditions and parameters). Then, the GMM estimator \( \hat{b} \) is defined as

\[
\frac{1}{I} \sum_{i=1}^{I} g(z_i, \hat{b}) = 0 \tag{25}
\]

It can be shown that this estimator is consistent:

\[
\hat{b} \rightarrow \tilde{b} \quad \text{almost surely as } I \rightarrow \infty
\]

and, under some technical assumptions, it has an asymptotic distribution

\[
\sqrt{I} \left( \hat{b} - \tilde{b} \right) \rightarrow N(0, \Sigma_b) \quad \text{in distribution as } I \rightarrow \infty \tag{26}
\]

where

\[
\Sigma_b = \left[ E \left( \frac{\partial g(z_i, \tilde{b})}{\partial b'} \right) \right]^{-1} S_w \left[ E \left( \frac{\partial g'(z_i, \tilde{b})}{\partial b} \right) \right]^{-1}
\]

where, if we assume additionally that

\[
E(g(z_i, \tilde{b}) g(z_j, \tilde{b}')) = 0 \quad \text{for all } j \neq i \tag{27}
\]

the \( S_w \) matrix is given by

\[
S_w = E(g(z_i, \tilde{b}) g(z_i, \tilde{b}')) \tag{28}
\]

To cast 2SNLLS into the GMM framework note that 2SNLLS can be seen as a GMM estimator that satisfies (25) for the \( g' = (g'_1, g'_2) \) function given by

\[
g_1(z_i, \delta, \gamma, \phi) = \begin{pmatrix} x_i \\ (x'_i, n_i) \phi \\ (x'_i, n_i) \phi \ r_i \end{pmatrix} \left[ H_i - x'_i \delta - (x'_i, n_i) \phi \gamma_1 - (x'_i, n_i) \phi \ r_i \ \gamma_2 \right]
\]

\[
g_2(z_i, \delta, \gamma, \phi) = \begin{pmatrix} x_i \\ n_i \end{pmatrix} \left( d_i - (x'_i, n_i) \phi \right)
\]

To prove consistency we have to check that this \( g \) function satisfies (24). That \( E \left[ g_2(z_i, \tilde{\delta}, \tilde{\gamma}, \tilde{\phi}) \right] = 0 \).
is trivial from (20). Using (20) to substitute for \((x', n_i)\alpha_0\) in the first set of expectations gives

\[
E \left[ g_1(z_i, \tilde{\delta}, \tilde{\gamma}, \tilde{\phi}) \right] = E \left[ \left( \begin{array}{c} x_i \\ (x', n_i)\phi \\ (x', n_i)\phi \cdot r_i \end{array} \right) \right] \left[ \varepsilon_i + u_i\tilde{\gamma}_1 + u_ir_i\tilde{\gamma}_2 \right] = A E \left[ \left( \begin{array}{c} x_i \\ n_i \\ (x', n_i) r_i \end{array} \right) \right] \left[ \varepsilon_i + u_i\tilde{\gamma}_1 + u_ir_i\tilde{\gamma}_2 \right]
\]

for some constant matrix \(A\) that consists of zeros, ones, and \(\tilde{\phi}\)'s. So we need to show that the expectation multiplying \(A\) is equal to zero.

By the above assumptions we have \(E(x_i\varepsilon_i) = E(x_iu_i) = E(n_i\varepsilon_i) = E(n_iu_i) = 0\). Note that (21) implies that \(u\) is uncorrelated with \(x\) and \(n\); if we strengthen this assumption to require that \(u\) is independent of \((x, n)\) then we have

\[
E \left[ \left( \begin{array}{c} x_i \\ n_i \end{array} \right) u_ir_i \right] = E \left[ \left( \begin{array}{c} x_i \\ n_i \end{array} \right) r_i \right] E(u_i) = 0
\]

where the independence assumption buys the first equality. Similarly, if we slightly strengthen assumption (19) to require that \(\varepsilon\) is independent of \(x\) and \(n\) then we have \(E \left[ (x', n_i) r_i [\varepsilon_i + u_i\gamma_1 + u_ir_i\gamma_2] \right] = 0\).

So, (24) is satisfied if we strengthen slightly the orthogonality assumptions and assume independence of \((\varepsilon, u)\) and \((x, n)\). Under this assumption the expression in (29) is zero so that 2SNLLS is consistent.

The asymptotic distribution can be found using (26). Here are the details. We need to estimate the matrix \(\mathcal{B}\)

\[
\mathcal{B} \equiv E \left( \frac{\partial g(z_i, \tilde{\delta}, \tilde{\gamma}, \tilde{\phi})}{\partial (\delta', \gamma', \phi')} \right)
\]

Taking derivatives of this \(g\) for each parameter we have:

\[
E \left( \frac{\partial g_1(z_i, \delta, \gamma, \phi)}{\partial (\delta', \gamma')} \right) = -E \left[ \left( \begin{array}{c} x_i \\ (x', n_i)\phi \\ (x', n_i)\phi \cdot r_i \end{array} \right) \right] \left( \begin{array}{c} x_i \\ (x', n_i)\phi \\ (x', n_i)\phi \cdot r_i \end{array} \right)
\]

38
\[
E \left( \frac{\partial g_2(z_i, \delta, \gamma, \phi)}{\partial (\delta', \gamma')} \right) = 0
\]
\[
E \left( \frac{\partial g_1(z_i, \delta, \gamma, \phi)}{\partial \phi'} \right) = E \left[ \begin{pmatrix} 0 \\ (x', n_i) r_i \\ (x', n_i) \phi \end{pmatrix} \right] \left[ \varepsilon_i + u_i(\gamma_1 + r_i \gamma_2) \right] - \begin{pmatrix} x_i \\ (x', n_i) \phi \\ (x', n_i) \phi r_i \end{pmatrix} (x', n_i) (\gamma_1 + r_i \gamma_2)
\]
\[
E \left( \frac{\partial g_2(z_i, \delta, \gamma, \phi)}{\partial \phi'} \right) = E \left[ \begin{pmatrix} x_i \\ n_i \\ (x', n_i) \right] (x', n_i)
\]

So, the derivative has the structure
\[
\mathcal{B} = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}
\]

Now, this derivative \( \mathcal{B} \) has to be consistently estimated. As is well known, one can obtain a consistent estimator \( \hat{\mathcal{B}} \) by replacing \( \delta, \gamma, \phi \) by the values estimated with \( 2\text{SNLLS} \), replacing the true residuals by calculated residuals calculated with \( 2\text{SNLLS} \) and replacing the expectation by sample averages in the expression for \( \mathcal{B} \). This yields
\[
\hat{B}_{11} = \frac{1}{I} (X, \hat{D}, \hat{DR})' (X, \hat{D}, \hat{DR})
\]
\[
\hat{B}_{12} = \frac{1}{I} \sum_{i=1}^{I} \begin{pmatrix} \varepsilon_i + \hat{u}_i(\gamma_1^{2\text{SNLLS}} + r_i \gamma_2^{2\text{SNLLS}}) \\ (x', n_i) r_i \\ (x', n_i) \phi \end{pmatrix} - \begin{pmatrix} x_i \\ (x', n_i) \phi r_i \end{pmatrix} (x', n_i) (\gamma_1^{2\text{SNLLS}} + r_i \gamma_2^{2\text{SNLLS}})
\]
\[
\hat{B}_{22} = \frac{1}{I} (X, N)' (X, N)
\]

Finally we need an estimate of \( S_w \). If we additionally assume that \( \varepsilon, u \) are independent across \( i \)'s this gives (27) then \( S_w \) is given by (28). Furthermore,
\[
S_{w,22} = E \left[ \begin{pmatrix} x_i \\ n_i \end{pmatrix} \begin{pmatrix} x_i \\ n_i \end{pmatrix} \right]' u_i^2 = E \left[ \begin{pmatrix} x_i \\ n_i \end{pmatrix} \begin{pmatrix} x_i \\ n_i \end{pmatrix} \right]' E (u_i^2)
\]
where the second equality follows from the independence assumption so that we can use the estimator

\[ \hat{S}_{w,2,2} = \frac{1}{I} \sum_{i=1}^{I} \begin{pmatrix} x_i \\ n_i \end{pmatrix}' \left[ \hat{\varepsilon}_i + \hat{u}_i \left( \gamma_1^{2SNLLS} + r_i \gamma_2^{2SNLLS} \right) \right]^2 \]

This will imply (as it should) the standard OLS var-cov matrix for the parameters of the "first stage regression" \( \hat{\phi} \).

The estimators for the other terms of \( S_w \) are

\[ \hat{S}_{w,1,1} = \frac{1}{I} \sum_{i=1}^{I} \begin{pmatrix} x_i \\ (x'_i, n_i) \phi^{OLS} r_i \end{pmatrix}' \begin{pmatrix} x_i \\ (x'_i, n_i) \phi^{OLS} r_i \end{pmatrix} \left[ \hat{\varepsilon}_i + \hat{u}_i \left( \gamma_1^{2SNLLS} + r_i \gamma_2^{2SNLLS} \right) \right]^2 \]

\[ \hat{S}_{w,1,2} = \frac{1}{I} \sum_{i=1}^{I} \begin{pmatrix} \gamma_2^{2SNLLS} + r_i \gamma_2^{2SNLLS} \end{pmatrix} \]

\[ \hat{S}_{w,2,1} = \hat{S}_{w,1,2} \]

Finally, a consistent estimator of the variance covariance matrix of 2SNLLS is:

\[ \hat{\Sigma}_{\delta, \gamma, \phi} = \hat{B}^{-1} \hat{S}_w \hat{B}^{-1}' \]

The reason for this complicated form is that it captures the fact that the var-cov of \( \delta, \gamma \) depends on the var-cov of \( \phi \), as it should. Because of the lower triangular form of \( \hat{B} \) this var-cov matrix for \( \hat{\Sigma}_{\delta, \gamma} \) depends on the var-cov matrix of \( \phi^{OLS} \). As is intuitive, the higher the uncertainty about \( \phi \) the higher the uncertainty about \( \delta, \gamma \). This variance covariance matrix can be used to build asymptotically valid hypothesis tests.
Bibliography


Figure 1: Impact of $\frac{c}{q}$ on $w_G$ and $NR(e_G)$
Figure 2: Impact of $k$ on $w_G$ and $NR(e_G)$
Figure 3: Impact of $\lambda$ on $w_G$ and $NR(e_G)$
Figure 4: Impact of $\alpha$ on $w_G$ and $NR(e_G)$
Figure 5: Impact of $A$ on $w_{\ell G}$ and $NR(\ell_G)$
Figure 6: Impact of $W/n$ on probability to win elections by incumbent.