Holding on? Ethnic divisions, political institutions and the duration of economic declines

Richard Bluhm and Kaj Thomsson*

This Version: December 18, 2019

Abstract

We analyze the duration of large economic declines and provide a theory of delayed recovery. We show theoretically that uncertain post-recovery incomes lead to a commitment problem which limits the possibility of cooperation in ethnically heterogeneous countries. Strong constraints on the executive solve this problem by reducing the uncertainty associated with cooperative behavior. We test the model using standard data on linguistic heterogeneity and detailed data on ethnic power configurations. Our findings support the central theoretical prediction: countries with more constrained political executives experience shorter economic declines. The effect is large in ethnically heterogeneous countries but virtually non-existent in homogeneous societies. Our main results are robust to a variety of perturbations regarding the estimation method, the estimation sample, measures of heterogeneity, and measures of institutions.

Keywords: economic crises, delayed recovery, political economy, ethnic diversity

JEL Classification: E61, O11, O43, J15, H12

*Bluhm *(corresponding author)*: University of California San Diego, Department of Political Science, and Leibniz University Hannover, Institute of Macroeconomics, e-mail: *bluhm@mak.uni-hannover.de*. Thomsson: Maastricht University, School of Business and Economics, e-mail: *k.thomsson@maastrichtuniversity.nl*

1 Introduction

Why are economic declines in Sub-Saharan Africa and some parts of the globe so persistent? In this paper, we propose a novel answer to this question which incorporates two well-known features of the subcontinent: high ethnic diversity and weak political institutions. We offer a theory of how these two interact during economic declines and test its empirical implications. Our main contribution is to outline a simple mechanism which links political heterogeneity and the powers of the executive to the repeated failure to agree on a policy response to an economic shock, even when the policy is economically effective and socially desirable. We explicitly focus on the process of policy formulation during the decline phase of a slump and consider the realities of countries with weak political institutions. This allows us to show that the uncertainty and distributional consequences created by imperfect constraints on the political executive lead to longer declines in ethnically heterogeneous countries.

Every crisis creates winners and losers. Our model highlights a commitment problem among those that benefit and those that suffer during the recovery process. *Ex ante* uncertainty about post-recovery incomes and a 'winner-take-all' effect caused by weak political institutions can lead to delays in the policy response. Delayed cooperation happens because ethno-political groups with political influence want to limit the risk of being expropriated, or exploited in some other way, by fortifying their own position (e.g. via strikes, demonstrations and riots, or moving resources out of harm's way). While we leave the precise characteristics of the policy response implicit, we assume that these groups are bargaining over some stabilization policy with between-group distributional consequences, such as a nationalization of a particular sector, cuts to food or fuel subsidies, or increases (decreases) in public investments.

We start from the premise that political power lies in the hands of the executive which distributes cabinet seats along ethnic lines. This broadly reflects the situation in Sub-Saharan Africa over the time frame under consideration (Francois et al., 2015) but we believe the theory captures how uncertainty and diversity interact during a crisis more generally. Outright expropriation of ethno-political groups is an extreme form of ethnic favoritism, which is often more subtle, but tends to decrease with stronger political institutions (e.g. Franck and Rainer, 2012, Hodler and Raschky, 2014, Burgess et al., 2015). However, such 'winner-take-all' effects, often linked to African politics, only operate at the margins in our model and can be offset by stronger constraints on the executive. Our theory is also not limited to non-democratic politics but can be understood in terms of ethnic parties in nascent democracies or, more broadly, in terms of diverse interest groups in established democracies with strong legislatures constraining the executive.

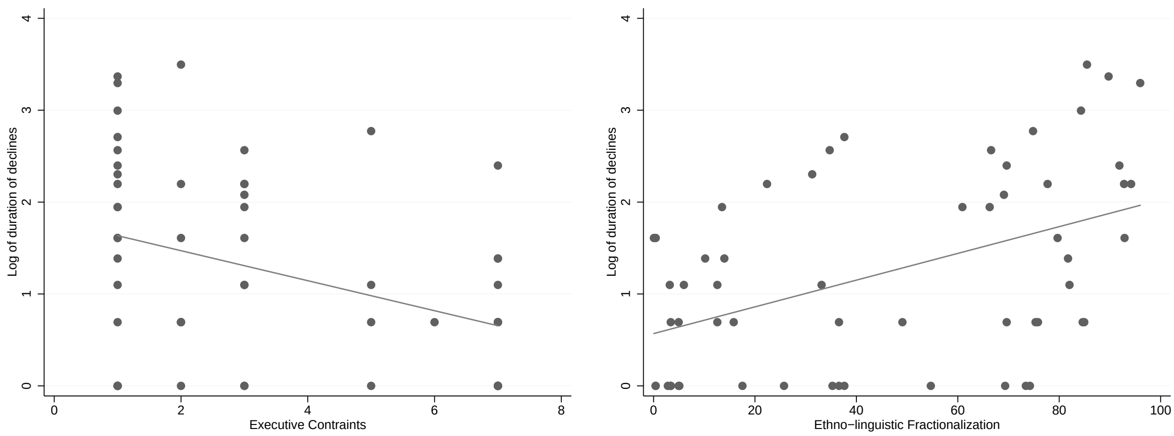
We derive three major insights from the model. First, delayed cooperation can occur in equilibrium, and weak constraints on the executive act as a political friction in ethnically diverse countries that can lead to social inefficiencies. Second, stronger political institutions can resolve this issue and bring about cooperation early on. Third, all else equal, the commitment problem between winners and losers is worse when the number of groups is large, and the introduction of institutional imperfections has a larger negative effect in more heterogeneous countries. We also outline an additional result which takes the relative size (strength) of the groups into account to show that political concentration matters for delayed cooperation. An important policy implication is that well-designed (and enforced) political institutions are key to containing the adversarial element of ethnic diversity and thus play a critical role in many developing countries.

Next, we take the model to the data. We first present a variety of partial correlations consistent with the proposed theory. We examine the central predictions using both standard data on linguistic heterogeneity and a more detailed data set which codes how much access ethnic groups have to the political executive. In line with our theory, we find that the effect of executive constraints on the length of declines is large in ethnically heterogeneous countries, but muted in ethnically homogeneous countries. This result is robust to many perturbations (e.g. different data sets, different measures and forms of heterogeneity, region and decade dummies, changing the functional form). We also show empirically that greater political concentration shortens declines and, *vice versa*, that a more even distribution of political power across groups increases delay. Possibly counter to intuition, we find that societies in which a larger group dominates or monopolizes the decision-making process experience shorter declines even when executive constraints are weak. In addition, we test a key assumption of the model and show that the number of ethno-political groups represented in the executive decreases during the early years of a recovery. Our empirical results are not exclusively driven by Sub-Saharan Africa even though the countries on the subcontinent are on average very diverse, have weaker executive constraints and tend to experience the longest declines.

Figure 1: Unconditional correlations with the duration of declines

(a) Executive constraints

(b) Ethno-linguistic fractionalization



Note(s): The figure plots the log of the duration of economic slumps over executive constraints in panel (a) and linguistic fractionalization in panel (b). The durations are based on the 58 slumps (in 51 countries) estimated on the Penn World Table 7.0 using the approach outlined in Bluhm et al. (2019). The duration of the decline phase is the time from the downbreak until the trough. No adjustment has been made for censored observations (unfinished declines). Constraints on the political executive are measured using an index scaled from 1 to 7 (least to most constrained) from the Polity IV data, and ethnic heterogeneity is proxied for by an index of ethno-linguistic fractionalization scaled from 0 to 100 (from Desmet et al., 2012).

Our dependent variable is the duration of economic declines. In line with our theory, we specifically focus on the duration of the decline phase of unexpected and large crises, whose beginning is defined by a negative structural break interrupting a previously positive growth trend. Such crises have large welfare consequences and can easily wipe out more than a quarter of GDP per capita in the course of several years (as in Mozambique, 1981–1986, or Zambia, 1968–2001). In fact, the duration of declines is five times longer in Sub-Saharan Africa than in Europe and twice as long as the world average. The methodology behind the econometric identification of these slumps is developed in a recent empirical contribution (Bluhm et al., 2019). That paper deliberately focuses on the duration of the decline phase for three reasons. First, the start of a crisis is often idiosyncratic and not necessarily related to a country's political institutions or level of social cohesion, but its duration depends on the relevant groups agreeing on coordinated responses. Second, the dynamics of recoveries differ a lot from the dynamics of declines (both empirically and theoretically). Third, most of the variation in the overall depth of slumps is due to the duration of the decline segment and not due to the rate of contraction. The motivating observation for this paper is that the duration (in years) until a recovery starts increases with greater ethnic divisions and decreases with stronger constraints on the executive. Figure 1 illustrates the unconditional correlation of the (log) duration of declines with executive constraints (-0.39) and ethno-linguistic fractionalization (0.47). Bluhm et al. (2019) are primarily concerned with the econometrics of identifying declines and establishing these stylized facts. The main objective of the current paper is to propose a theory that can explain these observations and to empirically examine our theoretical predictions using detailed data on ethnic groups and their access to political power.

Our work is also motivated by a literature in macroeconomics which emphasizes that economic growth is often not steady but instead characterized by different growth regimes. For example, it is well known that the correlation of growth rates across decades is low (Easterly et al., 1993). A key finding of the growth episodes literature is that growth accelerations are triggered by a variety of factors but are difficult to sustain (Hausmann et al., 2005, Berg et al., 2012). In developing countries, several years of positive growth can easily be followed by long and deep slumps. Such negative shocks can cancel out previous welfare gains and are often characterized by persistent output loss (Cerra and Saxena, 2008). This volatility starts to play a role in recent institutional theory, such as the 'limited access orders' of North et al. (2009). Little is known, however, about the deeper, more structural factors that are associated with longer (or shorter) slumps.

It is well established that ethnic heterogeneity is a fundamental determinant of economic prosperity. Heterogeneity is typically associated with low growth (Easterly and Levine, 1997), the undersupply of public goods (Alesina et al., 1999), and civil conflict (Fearon and Laitin, 2003, Esteban and Ray, 2011, Esteban et al., 2012). Ethnicity plays a rampant role in Sub-Saharan Africa where political organization is mostly ethnic, but ethnic diversity has also been linked to inadequate public good provision in US states (Alesina et al., 1999) or excessive deforestation in Indonesia (Alesina et al., 2014). Furthermore, high economic inequality among ethnic groups is associated with regional underdevelopment and political inequality (Alesina et al., 2016). Yet the role of ethnic heterogeneity during economic downturns has not been explored. Heterogeneity is, however, not necessarily a problem and is viewed favorably in many literatures. Furthermore, organizing along ethnic lines may resolve a contracting problem and help to enforce social sanctions within family or kin groups (Bates, 2000). In developed economies, the negative effects of ethnic heterogeneity may become muted, as skill complementarities matter more, or political institutions tame the conflict element inherent in diversity (Alesina and Ferrara, 2005). We formally incorporate this latter channel by showing that the negative effects of diversity on cooperative outcomes depend on the strength of political institutions.

The level of ethnic diversity is endogenous in the (very) long run. Heterogeneity is related to migratory distance from Africa (Ashraf and Galor, 2013), the duration of settlements and the history of the state (Ahlerup and Olsson, 2012), and variation in terrain and land endowments (Michalopoulos, 2012). At the micro-level, people may choose their group affiliation and switch groups depending on how discernible the individual features are which identify group membership (Caselli and Coleman, 2013). However, we do not expect ethnic compositions to change fundamentally in the short run (especially in the post-colonial period). That said, ethnicity is not always the most prominent political fault line in a society and the degree of access to political power of a particular group varies over time (Posner, 2004). The *Ethnic Power Relations* (EPR) data (Wimmer et al., 2009) codes the degree of access to executive power by different groups, focuses on politically relevant groups, and employs a more flexible notion of political division capturing the main fault line in a particular country (such as ethnicity, language, race or religion).[1] In our empirical part, we use these data to unpack the

[1] Francois et al. (2015) extend this approach further and provide data on the ethnic composition of the ministerial level in 15 African countries. If this data were available for more countries across the

model's predictions and contrast our findings to other established measures of diversity.

Ethnic diversity and weak political institutions often coincide. On the one hand, the adverse effects of ethnic heterogeneity may only be relevant in weakly institutionalized societies where political leaders often use (or abuse) ethnic and other divisions in their favor (Eifert et al., 2010). On the other hand, diversity affects the (endogenous) choice of institutions governing the executive power of such leaders (Aghion et al., 2004). There is some empirical evidence consistent with the view that ethnicity and political institutions interact.[2] However, the precise mechanisms behind how these two jointly determine the length of crises have not been investigated and may explain *substantial* parts of the robust negative correlation between ethnicity and growth. While plenty of anecdotal evidence exists, we are only aware of a paper by Rodrik (1999) which explicitly considers the link between ethnicity and negative growth.

The stylized facts motivating this paper cannot be explained by established theoretical frameworks. Ethnic groups could be engaged in 'wars of attrition' over the burden of reform, so that groups are trying to shift the costs of, say, a debt consolidation onto competing groups (Alesina and Drazen, 1991, Drazen and Grilli, 1993, Spolaore, 2004). Alternatively, we may view a slump through the lens of the 'veto player' literature, which suggests that more groups with blocking power will increase the time until a necessary reform is adopted (Tsebelis, 2002, Hicken et al., 2005, Gehlbach and Malesky, 2010). Both of these perspectives do not fully capture the scenario emphasized in this paper, where an interaction between weak institutions and group diversity is the key element generating delay. These alternate frameworks also deliver empirical predictions running counter to the evidence presented here.

The rest of the paper is organized as follows. In Section 2, we outline our model of how ethnic fractionalization and weak constraints on the executive can lead to delayed cooperation. In Section 3, we discuss the data, the empirical strategy and the main empirical results. Section 4 concludes.

globe, then it would be an ideal supplement to the EPR data used here.

[2] Collier (2000), for example, argues that ethnicity plays no role in democracies but reduces growth in autocracies and provides evidence along these lines. Easterly (2001) empirically investigates an interaction effect between institutions and ethnicity in determining growth and conflict.

2 Theory

We model group interactions during a slump as a cooperation game where groups decide on whether to formulate a policy response to a crisis that will initiate a recovery. We first focus on the symmetric two group case and then extend the model to allow for asymmetries and multiple groups.

2.1 Basic setup

We consider a population normalized to unity and split into J equal-sized (ethnic) groups. These J groups constitute the players of the game. Time is discrete and there is an infinite number of periods, indexed by $t = 1, 2, \ldots, \infty$. The per-period discount rate is δ. With slight abuse of notation, groups are indexed by $j = 1, 2, \ldots, J$, where $J = 2$ for the baseline model considered in this subsection. Each group acts as a single agent and we do not analyze internal coordination issues among members of the same group.

Preferences. Group j receives a net income of y_j in period zero. Total initial income in the economy is normalized to unity ($\sum_j y_j = 1$). Utility in each period is $g(y_j)$, where $g(\cdot)$ is increasing in y_j, concave and identical for all groups.

Slumps: decline and recovery. When a slump occurs, output declines by a fixed amount (Δ) in the first period. The income shock affects both groups proportionally and output remains at that level until both groups cooperate. Total income is now $1 - \Delta$ as long as the slumps lasts. Once a decision to cooperate has been reached, we assume that the economy recovers within one period. Groups decide to cooperate or not based on their expected future returns to cooperation.

We leave open the exact nature of the actions that can be taken to facilitate recovery. One example would be the implementation of a stimulus package in an economy well below potential output, possibly involving conditional loans from international financial institutions. Another possibility would be a bailout or nationalization of a banking sector at the verge of collapse or a bailout (nationalization) of a commodity sector accounting for a non-trivial share of the economy. There are plenty of other policy responses which will depend on the nature of the crisis. In developing countries with limited fiscal space, crises often coincide with cuts to government programs. Countries without floating exchange rates may be forced to devalue and/or impose capital controls. IMF or World Bank programs are often associated with cuts to fuel or food subsidies in developing countries and large adjustments to the public sector. Note that all of these policies are likely to have implications for the economic and political power of the affected groups. We explore a subset of such policies and their relation to the duration of declines in the empirical part of this paper.

Slumps: uncertainty. We assume that groups are uncertain about their post-recovery incomes—their economic position and political power relative to each other may change after the slump is over. In the baseline model, $y_1 = y_2 = 1/2$. Each group experiences a random shock to its income, where the probability that a group falls below a "threshold of safety" and is expropriated by the other group is given by a process p_t (which is explained in greater detail below). The setting is symmetric in the sense that p_t also gives the probability of the second group falling below the threshold. For the first group, the shock has support $\nu_1 \in [-y_1, 1 - y_1]$. Let w_1 denote actual income after the shock, so that y_1 is now a counterfactual; similarly for y_2, w_2. This implies that a slump will hit the groups unequally after recovery, but *ex ante* neither group expects to be hit harder.

Political institutions. We interpret executive constraints as limits on how much one can group gain or lose relative to the other through expropriation, as is common practice in the literature (see, e.g., Besley and Persson, 2011a,b). The intuition is as follows. If a particular group has been sufficiently weakened by the slump, the now stronger group may be in a position to expropriate part or all of the weaker group's income and exclude it from the political process. If the executive is unconstrained there are no checks on this type of predatory behavior. A complete constraint on how much one group can extract means that no expropriation can occur. A partial constraint implies that expropriation occurs only when one group becomes too weak. The dominant ethnic group controls the executive and shares the spoils from expropriating the weaker group with its members.

To be more precise, we model political institutions by including thresholds in the random shock. This introduces a second source of uncertainty: boundary outcomes (expropriation) are realized only beyond these thresholds. Let the parameter $c \in [0, 1/2]$ represent the weakness of executive constraints, and let the set $\mathcal{A} = [c, 1 - c]$ be the political "safe zone" in which there is no expropriation. Once a group falls below c, its income is expropriated (pushed to zero) and the other group gains the remainder. Thus, $1 - 2c$ can be interpreted as the ability of one group to commit to not expropriating the other group; alternatively, $c = 1/2$ can be thought of as the total lack of constraints.

To fix ideas, we interpret the 'winner-take-all' event as political extinction of the weaker group, though it can be understood in a variety of ways. In non-democratic politics, the threshold mechanism symbolizes the potential of some ethnic groups to exclude other groups from the political process and capture the rents of those that have been excessively weakened by the slump. In the terms of Francois et al. (2015), they would lose their co-ethnic ministers and hence their ability to influence politics and obtain patronage. In fact, we later show empirically that ethnic groups in plurality rule governments appear to lose power in the aftermath of a slump. Alternatively, it may even represent physical extinction due to ethnic conflict.

In democratic politics, it captures how strongly legislative institutions and the judiciary constrain the powers of the executive. If these institutions are able to represent all relevant interests and resolve the political uncertainty about post-recovery outcomes, then the political friction we emphasize disappears. However, even in consolidated democracies, imperfect constraints according to our definition capture the existence of thresholds that allow minorities to participate in government (e.g. the filibuster rule used in the U.S. Senate as well as several state legislatures, or, perhaps more fittingly in terms of identity politics, the 10% electoral threshold used in Turkey's general elections[3]).

In less established democracies, ethnic groups are often represented by parties

[3] This rule often kept Kurdish minority parties, but also other established parties, out of parliament. In the 2002 national elections, 46.33% of all votes were cast for parties below the 10% threshold and hence not represented in parliament.

reflecting their interests (see, e.g., Lublin, 2017). When the salience of ethnicity is high (e.g. in Kenya), parties and ballot box coalitions are fluid. They are often formed briefly before an election and dissolved again before the next, so that the legislature is at best of secondary importance. Ethnic politics is also far from an African phenomenon, but prevalent in many other countries, ranging from India over the Central Asian states to Bolivia. More generally, strong central states with weak constraints on the executive ruling over diverse identity groups can be found across the globe (e.g. in China, Russia, Turkey or Brazil).

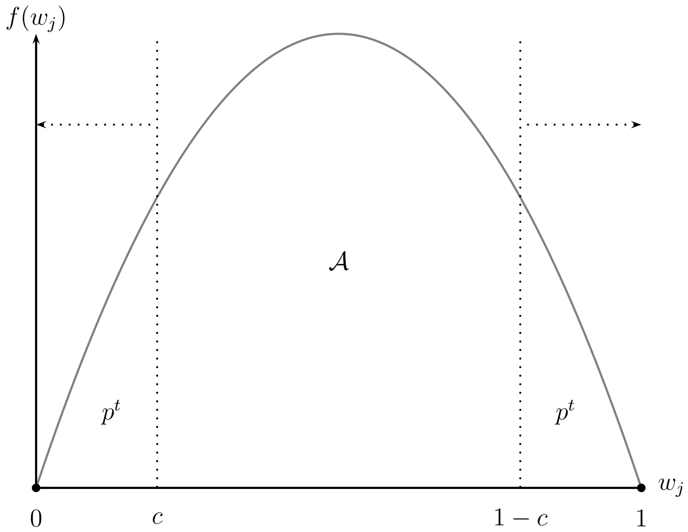
Delay. We assume that groups are able to fortify their position through non-cooperation. This is an essential assumption that will play an important role in the model. It implies that a group can (at least in part) counterbalance the uncertainty introduced by weak institutions through not cooperating, and thus potentially avoid falling below the threshold. The fortification of positions can be interpreted in multiple ways, with the appropriate interpretation depending on the context. In some countries, it can mean literal fortification, with political leaders mobilizing loyal members from an ethnically or regionally defined group for a show of force to the country's capital, or elsewhere to protect valuable resources. In societies less prone to violence and civil conflict, it can be interpreted as the building of political alliances, or the moving of production activities and resources to safe places (possibly outside the country).

Demonstrations and riots against cuts in food or fuel subsidies are commonplace in developing countries when the IMF or World Bank enter to stabilize a country (e.g. recently in Ecuador, Haiti, Egypt, Iraq, Sudan) while nationalizations of natural resource companies often occur during crises triggered by commodity shocks (e.g. strikes in the Bemba Copperbelt and the nationalization of Zambia's copper mines in the 1960s and 70s). Such actions limit the risk associated with a change in economic and political influence, but, most importantly, *they take time*. There are other interpretations of the gains from delay. Designing more sophisticated reforms in each period, for instance, decreases the risk of being inadvertently and strongly harmed by previous reforms.[4]

[4] We thank an anonymous referee for offering this suggestion.

In terms of our model, delaying cooperation limits the likelihood that a particular group will be expropriated. The parameter x is a measure of how much a group can reduce the risk of expropriation by holding out in each period. We assume that the probability of landing on either side outside the safe zone follows a linear process, so that $p^t = c - (t-1)x$ at each t when the groups can chose to cooperate or delay. Furthermore, we assume that expected utility conditional upon being in the safe zone is independent of p^t. Figure 2 gives an example of a distribution of w_j and illustrates the relevant regions.

Figure 2: Threshold effects as constraints on the executive



Note(s): The figure provides an example of a probability distribution of after-shock incomes $f(w_j)$ in the two group case. The weakness of executive constraints is given by the thresholds c and $1-c$. The safe-zone is the area denoted by $\mathcal{A}$. The probability of falling below or above the threshold at time t is denote p_t. A group which falls below the threshold c on the left will be expropriated so that $w = 0$ and the other (stronger) group ends up with $w = 1$ in return.

There is some conceptual overlap between our definition of executive constraints and the blocking power a group can exercise through non-cooperation. Constraints on the executive are institutionalized checks on the decision-making power of leaders. In our setting, perfect constraints resolve all uncertainty about experiencing unfavourable group-specific events and are undoubtedly associated with strong legislatures. In the absence of these formalized constraints, there is a *de facto* veto power arising from the ability to challenge particular policies "on the street" or contest who is governing the country. The recent literature on non-democratic politics argues that leaders are acutely aware of these constraints. North et al. (2009), for example, stress how countries first structure intra-elite relations before they can transition to more formal, democratic institutions.

The work by Bueno De Mesquita et al. (2003) models an incumbent's survival as a function of the size of the winning (or ruling) coalition and the selectorate, who expects to benefit in some way. Svolik (2009) presents a theory of contested dictatorships where power sharing with a ruling coalition arises from a revolution constraint. Building on this line of work, Francois et al. (2015) provide a model of non-legislative incentives and systematic evidence showing that such anticipatory power-sharing takes place in African politics. These are the broader, non-institutional constraints which are captured by the ability to hold out.

Timing. The following timing summarizes the structure of the game. At $t = 0$, the economy is in its initial state. Output $\sum_j y_j = 1$ is produced and shared equally.

1. At $t = 1$, the slump occurs, and incomes decline to $(1 - \Delta)y_j$. Both groups simultaneously choose to cooperate C or delay D.

2. For all $t > 1$, incomes remain at $(1 - \Delta)y_j$ if both groups did not cooperate in the previous period. They once again simultaneously choose whether to cooperate C or delay D. If, instead, there was cooperation in the previous period, incomes recover within one period, but are subject to a random shock and groups can land outside the political safe zone with twice the probability p^t. After a recovery, each group receives the same payoff as in the first post-recovery period forever.

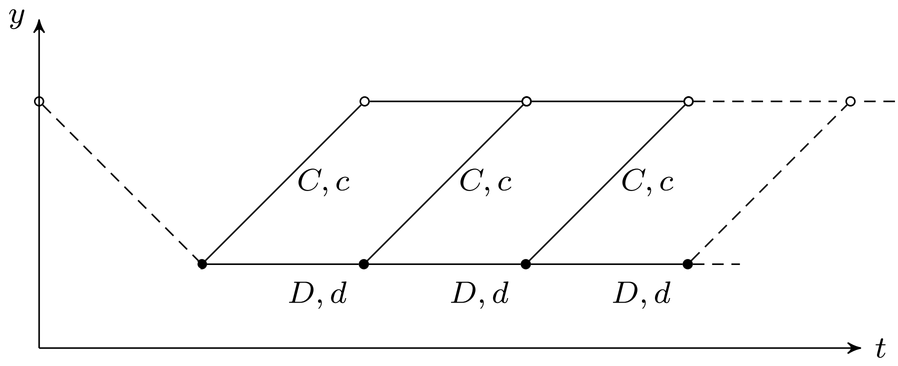
The present discounted value of the lifetime utility for each group is

$$v_j = \sum_{\tau=1}^{\infty} \delta^{\tau-1}\mathbb{E}g(\cdot) \tag{1}$$

where $g(\cdot)$ is $g((1 - \Delta)y_j)$ if the recovery has not yet occurred and $(1 - 2c)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + c(g(0) + g(1))$ otherwise. The discounted utility has two components: 1) if the economy has not recovered, groups are on a delay path, and 2) once the slump is over, they remain on a post-recovery path.

Figure 3 sketches how the economy evolves over time given different choices and presents a stylized view of the process we envision. Note that the action pair (D, d) has the same implication as (C, d) and (D, c); that is, cooperation of both groups is required for a recovery to occur.

Figure 3: A sketch of decisions and timing



Note(s): The figure shows a pseudo-game tree where the time path of each groups income (y) is displayed along side the solid choice nodes. When both groups cooperate, a recovery takes place and the same income is received forever on the post-recovery path. The level of this post-recovery income can differ for each group. When one or both groups choose delay, then the recession utility is received for another period, after which both groups will choose again.

The game has a symmetric structure. At each choice node (solid nodes), the comparison between any two adjacent periods always looks alike. The utility from cooperating in a particular period t when the other group cooperates in period t is

$$v_j^t(C, c) = \frac{1}{1-\delta}\left\{(1-2p^t)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^t(g(0)+g(1))\right\}. \tag{2}$$

Furthermore, assume the other group's strategy implies that it will cooperate in period $t+1$ as well, should the game reach that period. While cooperating in period $t+1$ is not necessarily optimal (after non-cooperation in period $t+1$), it it will be helpful to write down the utility from such an outcome. Specifically, the utility in period t from not cooperating in this period but cooperating in the following period $(t+1)$ when the other player's strategy implies that it would cooperate in both periods, is

$$v_j^t(D, c) = g((1-\Delta)y_j) + \frac{\delta}{1-\delta}\left\{(1-2p^{t+1})\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^{t+1}(g(0)+g(1))\right\}. \tag{3}$$

Although analytically trivial, it is useful for the coming analysis to establish the social optimum, before we characterize the non-cooperative equilibrium. Two points are worth highlighting. *i)* The utilitarian welfare-maximizing outcome involves no delay. This follows directly from the concavity of the utility function, as the sum of the group's utilities is maximized when their share is equal. At equal shares, the total welfare from any non-delay path dominates any delay path. *ii)* Any outcome with delay is Pareto dominated by some outcome without delay. To see why this is the case, take any path with delay, give the groups the same shares in every period, but let the recovery happen immediately. In this case, all groups receive more in the period before the recovery than they did with delay, and the same in every period after the recovery. The non-delay benchmark is particularly interesting when contrasted to the non-cooperative equilibrium of the game, where groups face a trade-off between immediately recovering and falling below the threshold c, or recovering later and reducing future uncertainty.

By comparing the utilities from cooperating in the first period and in the second period it is relatively straightforward to show that delay can occur in equilibrium. Our first result establishes this.

Proposition 1. *There exist parameter values, such that all equilibria involve delay.*

Proof. See Appendix. ■

The proof to the proposition shows that all components that make the immediate cooperation scenario less attractive are conducive to delay. The key issue rendering the cooperative equilibrium inaccessible is the *ex ante* commitment problem among potential winners and losers. Hence, worse institutions, or less ability to commit to not expropriate the loser (larger c), larger gains from holding out (larger x) and a larger value placed on the future (higher δ) make immediate cooperation less likely. Conversely, a larger shock (Δ) makes cooperation more attractive since a (potentially sizable) one period loss is avoided. The concavity of $g(\cdot)$ matters in the sense that it implicitly captures how averse groups are to negative events (falling below c) or how much they value expropriating other groups (landing above c). Without this risk aversion there would be no delay.

Note that the proposition is formally true only in a weak sense; it does not rule out that equilibria with immediate recovery could exist for some parameter values.[5] Rather, what the proposition establishes is that for some parameter values *all* equilibria are inefficient and welfare-suboptimal.

While still in the two group case, we can already highlight an interesting comparison to the homogeneous (one group) case.

Comment 1. *Without heterogeneity, there always exists an equilibrium with immediate recovery.*

Note that if the groups were to pool their resources as one, then all the elements inducing delay—except pure miscoordination—are absent. In other words, we need antagonistic political (ethnic) groups for the proposed mechanism to work, i.e. for the model developed here to provide a theory of why there is delay. A more careful analysis of group asymmetries and multiple groups follows in the model extensions.

To better understand when we are likely to see delay, we now characterize the subgame perfect equilibrium with (the earliest possible) recovery, if such an equilibrium exists. Given the symmetric structure of the game an interior solution exits and the optimal time to recovery can be derived using equations eq. (2) and eq. (3). Our second result summarizes a central insight of the model.

Proposition 2. *Stronger constraints on the executive shorten the time to recovery.*

Proof. See Appendix. ■

The proof shows that the optimal time to recovery is

$$t^* = \frac{g((1-\Delta)y_j) - \mathbb{E}[g(w_j)|w_j \in \mathcal{A}]}{x\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - (g(0)+g(1))\}} + \frac{c}{x} + \frac{1}{1-\delta} \tag{4}$$

[5] There are many "coordination failure" equilibria where neither group cooperates simply because they believe the other group will not. Such equilibria always exist, including an equilibrium with infinite delay. Our analysis, however, is focused on the more interesting scenarios (equilibria) where delay does not happen *only* as a result of this type coordination failure.

where the key comparative static result is $\partial t^*/\partial c > 0$.

This proposition says that if institutions are imperfect ($c > 0$), delay is going to be longer than if the groups are able to perfectly commit to not expropriating the losers.[6] In fact, the weaker the constraints on the executive (larger c), the longer is the expected time to cooperation. Intuitively, either group will find it optimal to delay until a point is reached when the benefit of holding out for an additional period is equal to the benefit of cooperating in this period, where the former may be the period in which all uncertainty regarding the political threshold is resolved. At this point, or the next discrete period, it is optimal to cooperate. Where exactly this point in time occurs depends on the trade off between recovering and potentially falling outside the political safe zone, or recovering later and reducing the remaining uncertainty.

For the remainder, we do not explicitly derive this equilibrium solution. Instead, we focus on the case where all uncertainty is resolved in the next period and compare different scenarios (e.g. perfect and imperfect institutions). We outline such an argument in the next comment.

Comment 2. *The existence of imperfect (weak) institutions makes delay more likely.*[7]

If institutions are perfect ($c = 0$), we have

$$\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] \geq g((1-\Delta)y_j) \tag{5}$$

and if institutions are imperfect (and $p^t = c$ for all t), we have

$$(1-2c)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + c(g(0) + g(1)) \geq g((1-\Delta)y_j). \tag{6}$$

Inequality (6) is harder to satisfy than inequality (5) under the concavity assumptions

[6] Again, this is only holds if we rule out equilibria involving immediate cooperation or infinite delay.

[7] Strictly speaking, a probabilistic statement (delay becomes "more likely") should not be used in this comment, as, for any given set of parameters, there either exists an equilibrium with immediate recovery or not. However, we follow Acemoglu and Robinson (2006) and say that a change in the setup of the model makes a particular outcome "more likely" if it becomes an equilibrium outcome for a larger parameter set.

imposed on the utility function. Note that this is entirely due to the presence of weak institutions ($c > 0$).

Discussion of the model. The baseline model focuses on several key aspects of the political economy of declines. First, we have modeled group interactions during crises under uncertain post-recovery incomes in a way that highlights that groups are not able to commit to compensating the losers. There are no enforceable contracts where the winners return the (additional) post-recovery gains, which is precisely the role played by strong constraints on the executive. Second, outcomes with delay can occur in equilibrium, and they do *not* coincide with the social optimum or with efficiency. Weak institutions act as a political friction creating potentially large economic inefficiencies. Third, heterogeneity matters and political groups are assumed to be willing to cooperate once it is optimal to do so. Entrenched distrust would only increase delay.

The model in this paper is developed specifically for understanding slumps (and their ending). In principle, one could imagine using a similar theoretical framework to predict the onset of other changes in the pace of growth, such as accelerations or extended periods of modest but successful development. However, growth accelerations, almost by definition, usually start from periods of unremarkable growth. They are unlikely to be preceded by the sense of urgency and extreme pressure to coordinate that provides the backdrop for the model we develop here. We also do not expect growth spurts to coincide with an elevated risk of expropriation, another key feature of our model, which we later document in the empirical section. With this in mind, we find it prudent to limit our claims of generality and present the model as a theory of economic declines.

We abstract from several features that would be potentially important in a paper with a different focus. For example, we assume the decline does not deepen after the first time period in a slump, and we assume that recovery, once it takes place, is immediate. Assuming an indefinitely continuing decline phase and non-immediate recovery would lead to a more realistic setup of the model. However, while these assumptions would add pressure to agree early on in a manner that might better reflect how slumps actually occur, the focus of our theory part is to understand the qualitative impact of heterogeneity and political institutions; we do not aim for quantitative predictions of the exact time to agreement. For similar reasons, we do not model the precise nature of the policy response, differentiate between democratic and autocratic regimes, or examine the impact of particular political constitutions (presidential or parliamentary). The exact form of the boundary events is also left open and could, for example, also represent the exclusion from public goods. We also do not differentiate between political and economic power. Again, such specificities are not essential to the main argument. Leaving them out does, however, imply that our paper might be a better description of some countries than others. The model is likely to be most relevant for understanding declines in countries where political divisions run along ethnic lines and executive power is shared.

Alternative explanations. The two leading alternative explanations for policy delay in a broad sense are 'war of attrition' models in economics and 'veto player' models in political science. Groups engaged in a war of attrition learn about the capacity of their opponents to bear the costs of waiting as time passes and stabilization occurs only once one of the groups concedes (Alesina and Drazen, 1991). A related class of models shows that a socially optimal reform may not be undertaken at all because it is *ex ante* not known to which (ethnic or other political) groups the benefits will accrue (Fernandez and Rodrik, 1991). Such models can also generate delay and an endogenous economic deterioration (e.g. Labán and Sturzenegger, 1994). Two key elements are shared with our approach: *i)* uncertainty about the expected outcomes, and *ii)* an *ex ante* commitment problem between (*ex post*) beneficiaries and losers of the reform. However, these models are not ideally suited for the setting we consider here, where a crisis is immediately obvious and the pre-crisis political power of each key player is often well known. Furthermore, these papers do not explicitly focus on ethnic diversity and presuppose the existence of strong political institutions. The empirical content also differs substantially from ours. For example, Drazen and Grilli (1993) stress that crises help stabilizations and Spolaore (2004) shows that political systems with a strong government (less constrained executive) reform more quickly. While our model is not directly comparable, our paper fundamentally differs in that crises coinciding with an unconstrained executive are at the heart of the problem.

The veto player framework in political science considers groups with the power to block changes to the status quo as obstacles to reform (Tsebelis, 2002). Contributions based on this framework generally find that policy stability is greater the more numerous the players in the political system that are required to agree to change. Veto player arguments have been used to explain why governments may agree on reforms necessary to combat an exchange rate crisis (e.g. Cox and McCubbins, 1997). Yet, the they have also been used to show the exact opposite. Gehlbach and Malesky (2010) find that a greater number of veto players weakens the influence special interests and can make reform more likely. Similarly, MacIntyre (2003) emphasizes a credibility gain from having fewer players commit to a policy following a currency shock. More generally, Tommasi et al. (2014) find that an intertemporal perspective can turn the standard veto player prediction on its head. Our intertemporal theory instead emphasizes that an important factor mediating the relationship between identity groups and delay are the limits placed on the executive. Within the veto player literature, Hicken et al. (2005) come closest to our paper, as they also stress the role of political institutions. However, they conclude that greater checks on the executive do *not* aid a recovery while a larger accountability group does. This stands in direct contrast to our predictions and empirical findings.

There are other, more context specific, mechanisms which could explain why we observe longer declines in heterogeneous countries with weak checks on the executive. Imagine a kleptocratic strong-man ruling over a diverse country, such as Mobuto's reign over the Democratic Republic of Congo (Zaire) from 1965 to 1997. If executive power is unchecked then capturing state resources for private gain is particularly easy. Ethnic diversity implies that there are many other relevant groups who have to be bought off, raising the need to exploit state resources further. Indeed, Mobuto was infamous not only for accumulating enormous wealth and eliminating political opponents, but also for maintaining an elaborate network of patronage and rotating ministers from different ethnic groups or regions (see, e.g., van Reybrouck, 2014). These elements are reflected in our theory, mostly in the form of (many) politically relevant groups who are afraid of suffering economically and losing access to the executive. What is not accounted for, is that a deterioration of checks on the executive (or an increase of social cleavages) could be the reason for the start of a crisis and also explain its subsequent duration.[8] We return to this issue later in the empirical section but note here that we cannot corroborate such a pattern using our data.

2.2 Extensions: asymmetric and multigroup settings

We now briefly sketch two extensions. To extend the model to the asymmetric and J-group cases, we make the following simplifying assumptions. First, we restrict attention to the uncertainty associated with falling below the political threshold. Specifically, we assume that if a group falls within the political safe zone its share of total economic activity will be equal to its pre-recovery share. Second, we use a piecewise linear utility function, in particular:

$$g(y_j) = \begin{cases} y_j & \text{for} \quad y_j > 0 \\ z & \text{otherwise} \end{cases} \tag{7}$$

where $z < 0$. Furthermore, for the case when there are more than two groups, we assume that at most one group can fall outside the political safe zone. We now work with a more general (continuous) probability function, where we only assume that $dp_t(y_j)/dy_j < 0$ for any given level of the constraint c. Finally, for simplicity, our comparative statics will be done for the case where all uncertainty is resolved after one period of delay.

How do changes in political concentration affect the political equilibrium? Intuition may suggest that smaller groups are more afraid of falling out of the political safe zone, implying that greater asymmetry between groups increases the likelihood of delay. However, our theoretical result suggests that the effect of changes in political concentration can go either way. Several things change in the two group case if the share of an initially weaker group moves closer to an equal allocation, so that the size of the

[8] We are grateful to an anonymous referee pointing out this option to us.

previously more powerful group decreases in return.[9] On the one hand, the emboldened group faces a lower probability of being expropriated. In addition, the group also has to forgo more utility in the delay scenario. Both work in favor of cooperation. On the other hand, the group now has more to lose if it gets expropriated and is thus less likely to cooperate. Without imposing further restrictions, the overall direction of the effect is undetermined and depends on the parameter values. We consider this an empirical issue and return to it in the next section. The following result gives the condition that has to hold for greater symmetry to lead to more delay.

Proposition 3. *A decrease in (political) concentration makes delay more likely, if the following condition holds*

$$\Delta + \frac{1}{1-\delta}\left\{\frac{dp_1(y_j)}{dy_j}(z-y_j) - p_1(y_j)\right\} < 0. \tag{8}$$

Proof. See Appendix. ■

Using this condition, we can summarize the circumstances that determine the direction of this effect.

Comment 3. *A decrease in concentration is more likely to work in favor of delay, if the shock is smaller, the future is less heavily discounted, the negative consequence of falling outside the political safe zone is greater and the probability of that event is not very responsive to the weaker group's share.*

Up until this point, we assumed that there are only two groups deciding on whether to cooperate or not. The final proposition relaxes this constraint and highlights two key insights of the model with respect to group heterogeneity (assuming symmetric groups).

[9] We usually interpret 'group power' or 'group size' to be roughly equivalent to the share of total income when discussing the model. In the empirical part, we do not observe group incomes and instead rely on the population shares of groups classified as politically-relevant.

Proposition 4. *i) An increase in the number of groups makes delay more likely.*
ii) Introducing imperfect (weak) institutions is more likely to lead to delay if the number of groups is larger.

Proof. See Appendix. ■

Contrary to the more equivocal result in Proposition 3, a larger number of groups decreases the likelihood of cooperation. The proof of part *i)* shows that the condition for immediate cooperation (when all uncertainty is resolved in the next period) boils down to an inequality that decreases in J. The intuition behind this proposition is simple. As the number of groups increases, every group becomes poorer and thus more vulnerable during a slump. Simplifying the model helps to show that this is driven by the uncertainty arising from the lack of executive constraints (which we now implicitly define through $p_t(y_j)$). Part *ii)* then takes the multi-group extension back to the motivating question behind Proposition 1; that is, we again explore why (when) Pareto-inefficient delays can occur, but now with a focus on the interaction between institutions and heterogeneity. It extends the first part of the proposition, which shows that more groups make cooperation less likely, and adds that an introduction of imperfect institutions will be particularly problematic when the number of groups is large.

So far we did not explicitly consider political power or political relevance. Instead we assumed that all groups start from inside the political safe zone, matter equally for the decision to cooperate, and may only fall into political irrelevance as a consequence of the slump. Keeping the decision mechanism fixed, we now reflect on what this implies for different power (group) configurations. We do so with an eye to the concepts that we can empirically capture in the next section. With this in mind, we summarize the role of political relevance in the last comment as follows.

Comment 4. *More politically relevant groups make delay more likely, while politically irrelevant groups do not matter. Conversely, this implies that if a group dominates or monopolizes the decision making process, delay becomes less likely.*

Here, political relevance refers to being a party to the negotiations and thus being included in this model. Note that this separates the issue of political relevance (and the number of groups) from the strength of constraints on the executive. For example, there may be a dominant group which is unconstrained and thus poses a threat for smaller groups, or there may be a dominant group whose hands are tied so that it cannot expropriate smaller groups. Since these cases are distinct, we should still observe an independent effect for both the number of groups represented in the executive (or different qualitative assessment of their access to executive power), and constraints on the executive *per se*.

As a final point before we move on to the empirical part, we want to highlight that any (positive) effect of executive constraints requires ethnic heterogeneity in our theoretical framework. This follows already from the baseline model. Furthermore, the second part of Proposition 4 shows that, when starting from a welfare-maximizing benchmark, an institutional imperfection that introduces the possibility of delay is more likely to matter in heterogeneous societies. Empirically, we therefore expect that an interaction term between a measure of ethnic heterogeneity and an index of the strength of political institutions should have a *negative sign* (i.e., reduce the duration of the decline phase).

3 Empirical Strategy and Discussion

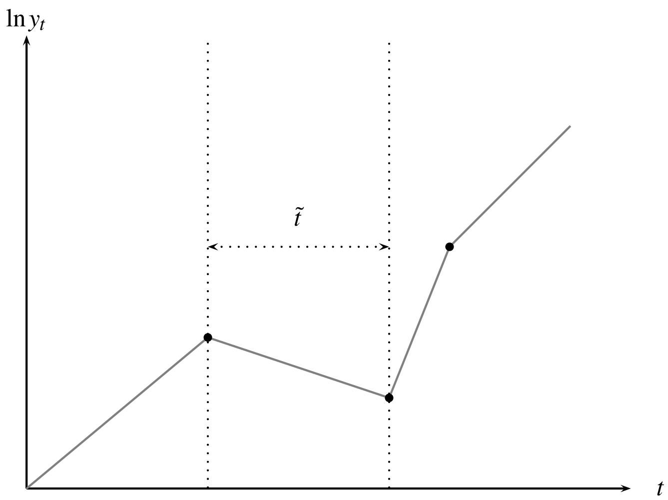
Decline spells. We characterize slumps by an abrupt negative departure from a previously positive growth regime that coincides with two successive trends breaks (usually separating a recovery regime from a post-slump regime). We then calculate the time it takes from the start of the first break until the empirical trough. Hence, our dependent variable is the duration of the decline segment during deep economic slumps. Figure 4 sketches the process we had in mind when designing this algorithm.

Our "to the bottom" definition of slumps differs from the existing literature which typically focuses on successive years of negative GDP growth, the duration from the business cycle peak to the trough, or the duration of entire recession (until full recovery).

However, this definition is perfectly in line with our theoretical set-up. It captures our argument that the length of the decline segment depends on the political system's ability to react to a crisis and accounts for the fact the dynamics of downturns and recoveries are often very different.

Structural break methods are an established way of identifying turning points towards negative or positive growth regimes (e.g. Rodrik, 1999, Hausmann et al., 2005, Jones and Olken, 2008, Berg et al., 2012, Papell and Prodan, 2014). They incorporate a notion of "pronounced" and "unexpected" declines in a univariate time-series sense, and allow us to statistically discriminate among multiple plausible starting points. However, since the identification of the duration of negative growth spells is not trivial and beyond the scope of this paper, we only briefly summarize the method here. More details can be found in the appendix of this paper and are discussed at length in Bluhm et al. (2019).

Figure 4: A stylized decline spell



Note(s): The figure shows a stylized example of the logarithmic time-series of GDP per capita before and after a major slump. The duration of the decline phase, denoted $\tilde{t}$, is the time (in years) between the first break and the through. The second break after the trough denotes another shift in the growth trend typically occurring after a recovery.

The procedure involves several steps. First, we fit a restricted partial structural change model with two breakpoints to each GDP per capita series. We impose sign restrictions on the model parameters, so that we only find major economic slumps. Second, we estimate candidates for the endogenous breakpoints and conduct a bootstrap Monte Carlo test of their significance. Third, we keep only breaks that are significant at the 10%-level and run the procedure again on the remaining data (before the first and after the second break) until all breaks have been found or the sample gets too small. Fourth, for each slump, we identify the empirical trough (the lowest point in the series after the beginning of the slump) and then compute the duration of the decline segment (denoted $\tilde{t}$). The spell is censored if pre-slump GDP per capita has not been recovered by the end of the sample, since we cannot rule out the possibility that the true trough occurs in the future.

Applying this algorithm to the *Penn World Table 7.0* yields 58 slumps in 51 countries from 1950 to 2008.[10] Table B-1 in the appendix provides a full list. Most slumps occur in the 1970s, 1980s and the early 1990s. The first observed slump begins in 1953 and the last observed slump begins in 1997. The method identifies many well-known slumps in both developing and developed countries, such as the Mexican debt crisis of the early 1980s (followed by the Tequila crisis in 1994) and the Finnish banking crisis of the early 1990s, but also deep and long-lasting declines in a number of African countries (e.g. Mozambique, 1981–1986, Togo, 1979–2008, or Zambia, 1968–1994). For the robustness checks, we also use a more lenient significance threshold of 20%, which results in a larger sample of 83 slumps in 70 countries and also includes many well-known episodes.

Table 1: Summary statistics of slumps

	Africa	Americas	Asia	Europe	Oceania	World
Countries	46	25	35	29	3	138
Countries with slumps	14	11	15	9	2	51
Number of slumps	14	16	16	9	3	58
Total years in decline	178	78	60	23	9	348
Duration of decline:						
– Min	1	1	1	1	2	1
– Median	16	2	2	1	3	3
– Mean	12.71	4.88	3.75	2.56	3.00	6.00
– Max	33	15	13	9	4	33
Incidence Rate	0.04	0.19	0.22	0.39	0.33	0.14

Note(s): The table shows summary statistics of the duration of the decline phase of slumps. A few countries (e.g. Chile) have repeated slumps which generates the discrepancy between the reported number of countries and number slumps. The incidence rate is defined as the number of exits from the decline period over the total years in decline.

Table 1 provides summary statistics of the slumps data. The basic correlations are as expected. Poorer countries have longer and deeper declines than richer countries;

[10] Note that we exclude countries with less than one million inhabitants and less than 20 years of data. We also use a more recent vintage with data until 2014 as a robustness check, see Table D-2.

countries in Africa have the longest and deepest spells. OECD countries do experience their fair share of volatility (12 slumps) but they tend to be shallow and short spells. The distribution of slumps is relatively even across the different regions, about 30-40% of all countries in each major region experience a slump. Their depth and average duration, however, varies greatly between developed and developing countries. We only observe a small number of repeated spells, three of which occur in Chile (starting in 1953, 1974, 1981). Ten slumps are unfinished; that is, GDP per capita has not recovered to the pre-slump level by the end of the period under investigation. Their trough is estimated to occur at the lowest observed value of GDP per capita and the spell is censored.

Measuring institutions. Our core measure of political institutions is the variable *Executive Constraints* from the Polity IV data set.[11] The variable directly measures the degree of institutionalized constraints placed on the political executive. It is coded unity when there is "unlimited executive authority" and seven when there is "executive parity or subordination"; intermediate values represent some constraints.[12] We believe that this variable corresponds well with the parameter c in our model. There are other indicators which directly measure the risk of expropriation, such as the property rights index of the Heritage foundation or similar indices by the International Country Risk Guide (ICRG). While they could in principle be used as a measure of c, these alternative series capture observed equilibrium outcomes, rather than the structural risk inherent in the political system (also see Acemoglu and Johnson, 2005, on this choice).[13] The Polity IV project has information on executive constraints annually from 1800 (or the year of

[11]An alternative measure of political constraints is the *POLCON* index proposed by Henisz (2000). This index is not our preferred measure for three reasons. First, the measure is derived from a veto-player model, while we propose a different theoretical approach. Second, it focuses on the number of parties in the legislature, not structural features of the executive. Third, it explicitly includes legislative fractionalization whereas we emphasize ethnic fractionalization of the executive. Additional results using this measure can be found in Table D-8.

[12]The variable explicitly considers the strength of the legislature, capturing the broad setting we consider in this paper, ranging from the de facto absence of meaningful parliaments over less established democracies with ethnic parties to consolidated democracies with independent legislatures. For example, one criterion for receiving the lowest score is "the legislature cannot initiate legislation or veto or suspend acts of the executive." A criterion for the highest score is "a legislature, ruling party, or council of nobles initiates much or most important legislation."

[13]There are also practical considerations leading us to prefer the Polity IV data. Our sample of breaks starts in 1950 and ends in 2008. The other data sources do not offer a consistent series covering the same period.

independence) until 2010. We do not use this wealth of time variation, since political institutions endogenously respond to the slump (Bluhm et al., 2019). We only rely on the degree of executive constraints in the last year *before* the slump and denote this variable $XCONST_0$. This rules out the possibility of feedback from the duration of declines to our measure of the risk of expropriation.

Measuring heterogeneity. We rely on two data sources to capture very different aspects of ethnic heterogeneity. The first source is a set of measures computed by Desmet et al. (2012) on the basis of the *Ethnologue* data. This data does not measure ethnicity directly but captures *linguistic* diversity. Fearon (2003) shows that linguistic (cultural) diversity coincides well with ethnic heterogeneity in some regions, notably Sub-Saharan Africa, but not so well in others. Together with the *Atlas Narodov Mira* data gathered by Soviet ethnographers in the 1960s, it is a standard source for data on ethnic heterogeneity and considerably more up-to-date than the former. Desmet et al. (2012) compute linguistic diversity at different levels of the language tree to capture the historical depth of ethnic divisions. We only make use of the most disaggregate level, since they also show that current divisions are correlated with economic growth more strongly than historical cleavages. The second data source is the *Ethnic Power Relations* (EPR) data presented in Wimmer et al. (2009). The EPR data has several advantages over other measures of linguistic or ethnic diversity, particularly for our application. It provides time series information on the degree of access to *executive* power of ethno-political groups from 1946 to 2010. Contrary to the *Ethnologue* data, it is not restricted to linguistic cleavages existing today. Instead, expert coders identify the most relevant division which may be ethnic, linguistic, racial or religious depending on the country and time period. The data contains information on the power status of each group, so that it allows us to focus on politically relevant groups; that is, groups with some form of representation in the presidency, cabinet, or other senior posts.

Our primary measure of heterogeneity is the commonly used index of ethno-linguistic fractionalization (e.g. Easterly and Levine, 1997). We consider it to be a summary metric

Table 2: Definitions of variables

Symbol	Description	Source and notes
	Dependent variable	
$\tilde{t}$	Duration of decline segment	Following Bluhm et al. (2019) computed using structural break model with a significance level of 10%. Underlying GDP per capita data is from the Penn World Table 7.0.
	Independent variables	
$XCONST_0$	Constraints on the executive	From Polity IV data. Measures *de facto* independence of the executive. Scaled from 1 (no constraints) to 7 (fully constrained). Fixed at last year before slump.
ELF	Ethno-linguistic fractionalization	From Desmet et al. (2012), the original source is the Ethnologue data (15^{th} edition). Cross-section.
ELF_0	Fractionalization of ethno-political groups	From Ethnic Power Relations data version 3.01 (Wimmer et al., 2009) supplemented with EPR-ETH 2.00 for small countries. Fixed at last year before slump.
POL	Ethno-linguistic polarization	From Desmet et al. (2012) using the Esteban and Ray (1994) measure with $\alpha = 1$ and $k = 4$. The original source is the Ethnologue data (15^{th} edition). Cross-section.
POL_0	Ethno-political polarization	Computed using Ethnic Power Relations data version 3.01 (Wimmer et al., 2009) and Esteban and Ray (1994) measure with $\alpha = 1$ and $k = 4$. Fixed at last year before slump.
ELA_0	Asymmetries between ethno-political groups (relative to fractionalization at equal sizes).	Computed using Ethnic Power Relations data version 3.01 (Wimmer et al., 2009) supplemented with EPR-ETH 2.00 for small countries. Fixed at last year before slump.
$GROUPS_0$	Number of groups	———
$EGIPGRPS_0$	Number of included groups	———
$EXCLGRPS_0$	Number of excluded groups	———
$DOMPOP_0$	Dominant population (in %)	———
$MONPOP_0$	Monopoly population (in %)	———
	Control variables	
$\ln GDPPC_0$	Log of initial real GDP per capita (*rgdpch*)	From the Penn World Table 7.0. Fixed at last year before slump.
Regional dummies	Africa, Americas, Asia, Europe, and Oceania.	UN classification. Oceania is base.
Decade dummies	1950s, 1960s, 1970s, 1980s, 1990s, and 2000s.	Coded at beginning of slump. 2000s is base.

for the presence of many relatively large groups. It is defined as

$$ELF_i \equiv 1 - \sum_{j=1}^{J_i} \left(\frac{n_{ij}}{N_i}\right)^2 = \sum_{j=1}^{J_i} \frac{n_{ij}}{N_i}\left(1 - \frac{n_{ij}}{N_i}\right) \tag{9}$$

where n_{ij}/N_i is the population share of group j in country i ($j = 1, 2, \ldots, J_i$, n_{ij} is the number of people in group j, and N_i the size of the population in country i). We employ two versions of this index: one computed by Desmet et al. (2012) and one computed using all group represented in the EPR data anchored to its pre-slump value (denoted ELF_0). We scale all heterogeneity indices by 100 to give changes on the right hand side a percentage point interpretation.

Another important dimension of diversity is the degree of polarization of a society. The literature on ethnic conflict often stresses that fractionalization and polarization have very different effects (e.g. see Esteban and Ray, 2011). We capture polarization with an index developed by Esteban and Ray (1994):

$$POL_i \equiv k \sum_{j=1}^{J_i} \left(\frac{n_{ij}}{N_i}\right)^{1+\alpha} \left(1 - \frac{n_{ij}}{N_i}\right) \tag{10}$$

where $\alpha = 1$ (as they show in an auxiliary theorem) and $k = 4$ to scale the index between zero and one. Again, we use a version computed by Desmet et al. (2012) and one we compute for the EPR data (denoted POL_0).

While the polarization index captures the extent of bimodality of a distribution[14], it is not a measure of asymmetries (such as the existence of one large and many small groups). To capture these, we propose a simple measure of ethno-linguistic asymmetries:

$$ELA_i \equiv \frac{\sum_{j=1}^{J_i} \left(\frac{n_{ij}}{N_i}\right)^2 - \frac{1}{J_i}}{1 - \frac{1}{J_i}} = \frac{J_i}{J_i - 1}\left[\sum_{j=1}^{J_i} \left(\frac{n_{ij}}{N_i}\right)^2 - \frac{1}{J_i}\right], \ \forall \ J_i > 1 \tag{11}$$

and $ELA_i = 1$ if $J_i = 1$. The ELA index is simply the (normalized) difference between fractionalization with equal shares and observed fractionalization; it's a normalized Herfindahl index. We only compute this index for the EPR data (denoted ELA_0).

[14] It attains its maximum at a symmetric bimodal distribution.

Desmet et al. (2012) do not report the data required for this measure. Recall that for any particular number of groups, the ELF measure attains its maximum at an equal allocation. The global maximum is reached when, in the limit, each person constitutes an ethnic group. Contrary to the polarization or fractionalization measure, the ELA index is zero when the groups are of equal sizes and approaches unity as a single group becomes dominant. For the empirical analysis that follows, using the index of group asymmetries together with the number of groups allows us to analyze the effect of these two components of ethnic heterogeneity separately and investigate the more subtle aspects of the theoretical model.

We also obtain several additional variables from the EPR data. $GROUPS_0$ is the number of relevant (active) ethno-political groups. $EGIPGRPS_0$ is the number of included ethno-political groups at the last year before the slump; that is, groups with have some level access to executive power. $EXCLGRPS_0$ is the number of ethno-political groups without access to the political executive. Finally, $DOMPOP_0$ and $MONPOP_0$ are the population shares of the dominant or monopoly groups (the two highest levels of political power occurring only in mono-ethnic governments). All of these variables are fixed at the last year before the slump to rule out any feedback from the duration to group composition. Table 2 describes all variables and lists the underlying data sources. Table C-1 in the Online Appendix presents the associated summary statistics.

Empirical approach. We employ standard event history techniques to study the duration of the decline phase. Survival analysis is particularly suitable for our purposes for two reasons. First, our empirical predictions are clearly about the time it takes until the recovery starts and parametric duration models allow us to fully specify the underlying duration process, including the shape of the baseline hazard. Second, survival methods are designed to deal with censored observations which account for a non-trivial proportion of our sample. If the observed slump is completed, then the likelihood incorporates the information that the recovery has started at some point within the sample. Whereas if the slump is unfinished, only the fact that the country was still experiencing a decline enters the likelihood. Our approach is to examine partial correlations and test whether these are consistent with the proposed theory. While we cannot rule out all forms of endogeneity, we do take care to ensure temporal precedence by only linking pre-slump realizations to the subsequent duration. Note that in a duration set-up with few repeated spells, we cannot use a full set of country and time effects (since time is already parametrized) but we do include region dummies and decade dummies in the robustness section.

To estimate the partial correlations, we run log-normal accelerated failure time (AFT) regressions of the form:

$$\ln \tilde{t} \equiv \ln(t - t_0) = \beta_0 + \beta_1 XCONST_0 + \beta_2 H + \beta_3 (XCONST_0 \times H) + \mathbf{x}_0' \boldsymbol{\xi} + \epsilon_t \tag{12}$$

where $\tilde{t}$ is analysis time, t_0 is the last year before the slump, $XCONST_0$ is executive constraints, H is a measure of group (ethnic) heterogeneity, $\mathbf{x}_0$ is a vector of controls, and $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon)$. Variables which could endogenously react to a prolonged duration of declines are kept fixed at t_0 to rule out any such feedback; if they have no time dimension, then we drop the subscript. All parameters are estimated using Maximum Likelihood and the standard errors are clustered on the country level to account for repeated spells.

Our main parameters of interest are β_1, β_2, and β_3. In several regressions, we impose $\beta_3 = 0$ to estimate first-order effects before examining the hypothesized interaction effect. The vector $\mathbf{x}_0$ may include additional heterogeneity measures, the (log of) initial GDP per capita before the slump, region effects and decade dummies.

Accelerated failure time models are so named due to their interpretation. A coefficient greater than zero implies that time passes more slowly, so that the exit of the decline phase is prolonged. A coefficient less than zero implies that time passes more quickly and hence that the recovery starts sooner. Alternatively, we may simply read the effects as elasticities (or semi-elasticities) of the expected duration with respect to the variables on the right hand side. Duration models have the main benefit of accounting for right censoring, otherwise their interpretation is identical to log-linear OLS when they are cast in the log-normal AFT form.

Results. Table 3 presents the first set of results corresponding mainly to the predictions derived from the baseline model. We compute two sets of estimates. One using the *Ethnologue* data which focuses on linguistic diversity and one based on the EPR data which incorporates only politically-relevant groups divided along the predominant social cleavage (ethnic, linguistic, racial, etc.).

Columns (1) to (3) use the *Ethnologue* data. Column (1) establishes that stronger constraints on the political executive shorten the expected duration of the decline phase and that greater linguistic heterogeneity has an adverse effect on the expected duration. The effects are statistically significant at the 1%-level and economically meaningful. A one point improvement in executive constraints (before the slump) leads to an approximate 17.6% reduction in the duration until the trough. Conversely, a one percentage point change towards greater linguistic heterogeneity prolongs the decline phase by about 1.7%. Column (2) allows for a conditional effect and strongly suggests that the effect of political institutions depends on the level of linguistic diversity (and *vice versa*). Whenever we introduce an interaction term, we first center the two constituent variables on their average. This shifts the coefficients of the two base levels into a meaningful range, but leaves the magnitude and statistical significance of the interaction term unaffected. Holding the other variable constant, the coefficient on either base variable now measures the effect of a one unit change away from the average. As a result, the interaction effect can be ignored; it has to be taken into account only when both variables change. The interaction between executive constraints and linguistic fractionalization is significant at the 5%-level and comparatively large. The specification predicts that at perfect homogeneity the median decline lasts about 2 years, while at perfect heterogeneity it lasts about 12 years. These estimates cover nearly all of the observed differences between declines in Western Europe and Sub-Saharan Africa.

The results in columns (1) and (2) are consistent with our theoretical predictions; greater constraints on the executive shorten the expected duration unless the society is nearly homogeneous. The partial effect of executive constraints is not statistically different from zero for low ELF values. Column (3) adds the linguistic polarization

Table 3: Baseline – Executive constraints, heterogeneity and interactions

	Dependent variable: $\ln \tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Ethnologue			*Ethnic Power Relations*		
$XCONST_0$	-0.193***	-0.289***	-0.175***	-0.187***	-0.262***	-0.170**
	(0.060)	(0.084)	(0.062)	(0.067)	(0.085)	(0.067)
ELF	0.017***	0.019***	0.023***			
	(0.004)	(0.004)	(0.006)			
$XCONST_0 \times ELF$		-0.004**				
		(0.002)				
POL			-0.011			
			(0.007)			
ELF_0				0.020***	0.023***	0.025***
				(0.007)	(0.007)	(0.007)
$XCONST_0 \times ELF_0$					-0.004*	
					(0.002)	
POL_0						0.012
						(0.009)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	48	48	48	47	47	47
Spells	58	58	58	57	57	57
Years of decline	348	348	348	346	346	346
Log-$\mathcal{L}$	-74.704	-72.495	-73.645	-76.294	-74.952	-75.597
Pseudo-R^2	0.150	0.175	0.162	0.119	0.134	0.127

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

measure to the specification in column (1). The literature on civil conflict stresses that polarization matters; e.g. Esteban and Ray (2011) show theoretically that conflict over public goods is driven by polarization and conflict over private goods by fractionalization. Contrary to this literature but in line with our model, we find no evidence in favor of the hypothesis that polarization is an issue for (the lack of) cooperation during declines, while the coefficient on fractionalization is robust to this perturbation. In other words, existence of two equally powerful groups does not predict longer declines than, say, three equally-sized linguistic groups.

Measures of linguistic diversity tend to describe Sub-Saharan Africa as more diverse in comparison to other regions than alternative diversity measures. This begs the question if we are just estimating an "Africa effect." Columns (4) to (6) use the EPR data which addresses this issue by alternating the relevant cleavage by country (from racial over linguistic to religious). This changes the relative location of Sub-Saharan Africa, which is only the second most diverse region on this measure, after South Asia, contrary to being the most linguistically diverse region based on the *Ethnologue* data. The EPR data also only codes politically relevant groups, so that the level of heterogeneity—no matter the measure—is generally lower. Note that we compute the heterogeneity measures for all politically relevant groups, not just the included groups. Strikingly, the results are virtually unchanged. Column (4) shows that the first order effects of executive constraints are the same, and the effect of fractionalization is well within one standard error of the previous estimate. Moreover, the sign and size of the interaction effect in column (5) is nearly identical to the one in column (2). Only the statistical significance of the interaction effect is a bit lower (cluster robust t-stat = -1.71). Column (5) shows that we also find no evidence in favor of *ethno-political* polarization affecting the duration of declines, just as with linguistic polarization. Contrasting these results to the *Ethnologue* data, it seems safe to conclude that we are not only explaining that declines in Sub-Saharan Africa last longer than elsewhere because the subcontinent is the most linguistically diverse. Our results also hold when we account for political relevance and vary the relevant divisions so that Sub-Saharan Africa is no longer the most diverse region in the world.

Overall, Table 3 provides significant evidence that there is a robust partial correlation of the duration of the decline phase with executive constraints on the one hand and with ethnic diversity on the other hand. In addition, the effect of weak constraints on the executive seems to be conditional on the degree of (ethno-political) fractionalization.

In Table 4 we "unpack" these statements further and examine what type of group configurations give rise to the cooperation problem we are analyzing. These results correspond mainly to the empirical content of the model extensions. We now only use the EPR data, as it provides the necessary detail on the number of groups, their power status and more. Column (1) presents a result that may appear puzzling at first sight. If we measure heterogeneity simply by the number of ethno-political groups, then we find no evidence of an effect on the duration of declines. However, this coefficient amalgamates two effects. The EPR data distinguishes between *included* groups, which have access to executive power, and *excluded* groups, which lack power at the state level or are (at worst) discriminated against. In the model presented earlier, only the former are relevant players and thus we also expect that only they are empirically relevant. Column (2)

Table 4: Extensions – Number of groups, political relevance, and asymmetries

	Dependent variable: $\ln \tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$XCONST_0$	-0.225***	-0.241***	-0.215***	-0.179***	-0.210***	-0.196***
	(0.070)	(0.063)	(0.065)	(0.066)	(0.070)	(0.066)
$GROUPS_0$	-0.008			-0.031**		
	(0.018)			(0.014)		
$EGIPGRPS_0$		0.426***			0.290**	0.285**
		(0.095)			(0.124)	(0.113)
$EXCLGRPS_0$		-0.012			-0.021*	-0.013
		(0.013)			(0.012)	(0.012)
$DOMPOP_0$			-0.007*			
			(0.004)			
$MONPOP_0$			-0.011**			
			(0.005)			
ELF_0				0.022***	0.013	
				(0.007)	(0.009)	
ELA_0						-0.012**
						(0.005)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	47	47	47	47	47	47
Spells	57	57	57	57	57	57
Years of decline	346	346	346	346	346	334
Log-$\mathcal{L}$	-81.069	-75.062	-77.647	-75.457	-73.253	-71.791
Pseudo-R^2	0.064	0.133	0.103	0.129	0.154	0.171

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

confirms this expectation. The effect of included groups is statistically significant at the

1%-level and economically very large: an additional group increases the duration of the decline phase by about 53%. On the contrary, the effect of excluded groups is estimated to be near zero and has a comparatively tight 99% confidence interval centered near zero. In line with the theory, these results suggest that only ethnic groups with some degree of access to political power matter for the duration of declines.

We have not yet isolated whether this adverse effect of heterogeneity is due to several equally powerful groups co-existing in the society or due to particular asymmetries in political power. Columns (3) to (6) represent different attempts towards empirically answering this question. In Comment 4, we translated the theoretical results regarding political concentration from Proposition 3 and the number of groups from Proposition 4 into the concepts of monopoly groups and dominant groups. Column (3) is the empirical counterpart. Here we relate the share of population represented by a group that either monopolizes or dominates the political executive to the duration of the decline phase. The results are unambiguous. Both variables are associated with substantially shorter declines. Through the lens of our framework, this finding is hardly surprising. The definition of dominant or monopoly groups means that they rule alone and are thus not bargaining over stabilization policies with other groups in the executive.

Columns (4) and (5) try to explicitly tackle the issue of the number of groups versus group asymmetries. In column (4), we include the number of ethno-politically relevant groups together with the index of ethno-political fractionalization. This leads to an interesting *ceteris paribus* condition. Increasing the degree of fractionalization by one percentage point while holding constant the number of groups necessarily implies that political concentration is decreasing; that is, the groups are becoming more alike. Recall that for any given number of groups, fractionalization is maximized at equal shares. The estimates thus suggest that less political concentration leads to longer declines. Column (5) again distinguishes between included and excluded groups to illustrate that only the former are relevant. The coefficient on the ELF_0 measure loses significance, suggesting that the number of included groups may drive the effect of ethnic heterogeneity and that group imbalances hardly matter. However, column (6) addresses this issue more directly by using our index of ethnic asymmetries and provides the same answer as column (4). Now the effect is easy to interpret, negative and significant at the 5% level. A one percentage point move towards greater asymmetries (political concentration) shortens the duration by about 1.2%. Note that the effect of executive constraints remains robust throughout, fluctuating around a 20% reduction in the duration of declines.

To summarize, Table 4 adds several valuable insights about the effect of ethnic diversity on the duration of declines. Fractionalization of linguistic or ethno-political groups masks two effects: 1) the expected duration is increasing in the number of politically relevant groups, and 2) the expected duration is decreasing in greater group asymmetries (political concentration). Both theory and evidence suggest that this is not an issue of polarization, but rather an issue of adding more politically-relevant groups.

Table 5 selects three key specifications using both data sources and then subjects them to two robustness checks. First, we return to the issue of whether we are estimating an "Africa effect" by including region dummies in each specification. Second, we control for temporal heterogeneity by including a dummy for the decade in which the slump began in every other specification, since the 1970s, 1980s and 1990s exhibit significantly higher volatility than the other decades.

The Africa dummy is significant in all specifications, capturing that declines take substantially longer on the African continent. Nevertheless, we still find comparable effects. Column (1) uses the *Ethnologue* data and shows that our two variables of interest are robustly correlated with within region differences in the duration of declines. Column (2) adds that this is still the case when we also control for temporal heterogeneity. Using the *EPR* data, columns (4) and (5) verify that the same holds for ethno-political fragmentation. The standard error of the interaction term becomes somewhat wider, leading to a loss of significance, but the estimated coefficient is extremely stable. The last two columns show that this also holds for the effect of the number of included groups. In general, there is significant evidence of regional heterogeneity (a χ^2-test always rejects the null of no heterogeneity at the 5%-level), but there is somewhat less evidence of temporal

Table 5: Robustness – Region and time effects

	Dependent variable: $\ln \tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$XCONST_0$	-0.256***	-0.251***	-0.211***	-0.171**	-0.181***	-0.145**
	(0.071)	(0.060)	(0.077)	(0.068)	(0.066)	(0.064)
ELF	0.020***	0.019***				
	(0.004)	(0.003)				
$XCONST_0 \times ELF$	-0.003***	-0.004***				
	(0.001)	(0.001)				
ELF_0			0.017**	0.015**		
			(0.007)	(0.006)		
$XCONST_0 \times ELF_0$			-0.003	-0.003*		
			(0.002)	(0.002)		
$EGIPGRPS_0$					0.298***	0.212*
					(0.097)	(0.119)
$EXCLGRPS_0$					0.024*	0.009
					(0.013)	(0.026)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Decade dummies	No	Yes	No	Yes	No	Yes
Summary stats						
Exits	48	48	47	47	47	47
Spells	58	58	57	57	57	57
Years of decline	348	348	346	346	346	346
Log-$\mathcal{L}$	-63.635	-58.134	-67.966	-64.108	-68.705	-66.701
Pseudo-R^2	0.276	0.338	0.215	0.260	0.207	0.230

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

heterogeneity (on top of duration dependence).[15] Throughout Table 5 the coefficient of political institutions and the coefficients of the various measures of ethnic heterogeneity remain statistically significant at conventional levels and well within their usual range.

We report a battery of robustness checks in the Online Appendix. Table D-1 uses a more lenient threshold for the identification of slumps (a significance level of 0.2). Our main results are qualitatively and quantitatively comparable in this larger set of 83 episodes. Table D-2 uses the *Penn World Table 9.0*, which includes data through 2014,

[15] A χ^2-test rejects the null of no temporal heterogeneity at the 1%-level in column (1), at the 5%-level but not the 1%-level in column (4), and fails to reject the null at conventional levels in column (6).

and shows that our findings findings do not depend on the vintage of the GDP data or the exclusion of the Great Recession starting in 2008. Table D-3 and Table D-4 change the resampling technique from a recursive parametric bootstrap to a fixed-design bootstrap for both the strict and more lenient thresholds. This substantially increases the number of slumps (up to 106 spells) but hardly affects our main conclusions. Table D-5 exchanges the fractionalization data with data on ethnic, linguistic and religious heterogeneity from Alesina et al. (2003), data on ethnic and cultural distance from Fearon (2003), and the original *Atlas Narodov Mira* data. For all but religious fractionalization, we find very similar interaction effects. In Table D-6, we include the Greenberg index, which is a Gini index that accounts for linguistic distance, a peripheral heterogeneity index, which is a variant of the Greenberg index that accounts for the alienation between groups in the center and the periphery (both are from Desmet et al., 2009), genetic diversity from Ashraf and Galor (2013), and three measures of segregation from Alesina and Zhuravskaya (2011). Again, all interactions other than the one with religious segregation point in the right direction, while measures of *linguistic* heterogeneity tend to have the most robust effects.[16] This is in line with the extant empirical literature which tends to find that religious diversity often plays a different role than ethno-linguistic diversity (e.g. see Alesina et al., 2003, Alesina and Zhuravskaya, 2011). Next we examine if spatial and ethnic inequalities capture relevant notions of heterogeneity. We use a cross-section of spatial and ethnic inequalities in 1992 from Alesina et al. (2016) who estimate inequality based on differences in nighttime light intensities among arbitrary boxes, *Ethnologue* homelands at different levels of the linguistic tree, and homelands from the *Atlas Narodov Mira*. Table D-7 shows that while the coefficients on the constituent variables and the interaction term consistently point in the right direction, there is only weak evidence that ethnic inequality among linguistically distant groups (level 1) leads longer declines and interacts with political institutions. Table D-8 switches the Polity IV data with the

[16]We also broadened our concept of political heterogeneity to government fractionalization and legislative fractionalization. We then ran horse race regressions of these variables and their interactions with executive constraints, while keeping ethno-linguistic fractionalization, GDP per capita, region dummies, and time dummies in the specification. None of these variables or their interaction with executive constraints are significant at conventional levels, while the effect of ethno-linguistic fractionalization remains robust throughout.

political constraints data from Henisz (2000). Here too, the main results remain intact for the alternate measures of executive constraints. Table D-9 illustrates that neither changing the periodicity of the initial period fixed effects nor the simultaneous inclusion of several heterogeneity measures alter our results. Table D-10 adds that these results do not depend on the specific functional form of the survival process. Finally, Table D-11 includes several policy variables (government size, changes in the exchange rate regime, and IMF or World Bank projects) to illustrate that *i)* endogenous policy choices seem to be related with a shortening of the decline period and *ii)* these effects remain but become somewhat weaker when we include our main variables of interest. We interpret this as additional evidence of having uncovered a deeper mechanism which only partially operates through these proxies for policy choices.

Next we test a key assumption of our model. Throughout the paper we assume that weaker groups face an elevated risk of loosing political influence during a crisis or, more specifically, during the recovery period. For this test, we turn things around and run group-level regressions of the probability that an ethnic group is currently part of the executive on the timing of the crisis. This implies that we are now dealing with a considerably larger country-group-year panel over the period from 1950 to 2008. To purge most of the confounding heterogeneity, we always include group fixed effects, time fixed effects and country-specific linear time trends in the specifications.

Table 6 demonstrates that the data are consistent with our approach. The table shows that *individual* groups with junior or senior partner status in multi-ethnic governments face a statistically significant and non-trivial probability of falling out of the executive in the five years after the through (but do *not* tend to fall out immediately during the decline phase). The implied effects are sizable as well. A group faces 5.5 percentage points increase in the risk of not sharing executive power in the first year after the trough. Table E-1 in the Online Appendix shows that these results hold with standard errors clustered on the country level. This adds another piece of evidence and suggests that our simplified focus on expropriation during recoveries has a basis in reality.

A remaining concern may be that the occurrence of slumps could driven by a sudden

Table 6: Number of politically-relevant groups

	Dependent variable: $Pr(EGIP)_{git}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Decline in t	0.004	0.002	0.000	-0.001	-0.002	-0.002
	(0.018)	(0.018)	(0.019)	(0.019)	(0.019)	(0.020)
Trough in $t-1$	-0.055***					
	(0.018)					
Trough in $t-2$ to $t-1$		-0.044**				
		(0.018)				
Trough in $t-3$ to $t-1$			-0.037**			
			(0.018)			
Trough in $t-4$ to $t-1$				-0.032*		
				(0.017)		
Trough in $t-5$ to $t-1$					-0.030*	
					(0.017)	
Trough in $t-6$ to $t-1$						-0.026
						(0.017)
Control sets						
Group FEs	Yes	Yes	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes	Yes	Yes
Country-specific trends	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Adjusted-R^2	0.136	0.137	0.137	0.137	0.137	0.136
$\bar{G}$	150	150	150	150	150	150
$\bar{T}$	36.67	36.67	36.67	36.67	36.67	36.67
$\bar{G} \times \bar{T}$	5500	5500	5500	5500	5500	5500

Note(s): The table shows the results from linear group-level panel regressions of the probability of being a politically relevant group with government participation on our indicators of the timing of economic slumps. The group-level data is from the EPR-ETH 2.00 (which includes small countries missing in EPR 3.01). The dependent variable is called `status_egip` in the original data. All groups with a power status of 'junior partner' or higher are coded as unity, all others as zero. We drop countries ruled by dominant and monopoly groups as these are single-ethnicity governments which are not the object of interest in our theory. The standard errors are clustered on the group level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

deterioration of political institutions which then also prolongs their duration. While we cannot rule out that this is true in some cases, a brief analysis of the timing of slumps strongly suggests that these two do not coincide systematically. Table F-1 and Table F-2 in the Online Appendix document the absence of any partial correlation between executive constraints (or an interaction with ethnic heterogeneity) and the onset of slumps. Instead, macroeconomic shocks, such as commodity terms of trade shocks and runaway inflation, are correlated with the timing of slumps. There are several cases supporting this narrative in our data. The Mexican crisis from 1982 to 1988, for example, is triggered by a decline in the oil price which rendered the government's strategy of borrowing against future oil revenue untenable and resulted in a sovereign default. Mexico is home to several indigenous groups but politics are dominated by the overwhelming majority of Mestizos. Constraints on the executive were at an intermediate level before the debt crisis. Similarly, Zambia's collapse from 1968 until 1994 occurs when world copper prices begin to enter a period of volatility and secular decline. Zambia is incredibly diverse, with more than 70 tribes, and few constraints were placed on the executive immediately after independence. Moreover, President Kenneth Kaunda was known for ethnically balancing his cabinet according to population size (Posner, 2005), just as assumed by our model. It thus comes at no surprise that this is one of the longest declines we observe in the data.

Taken together, these last empirical findings tell the following story. First, ethnic heterogeneity and constraints on the political executive are robust determinants of the length of the decline phase during economic slumps. Second, this result is not due to regional differences in ethnic heterogeneity but holds when we only use *within region* variation. Third, our main findings are robust to a variety of perturbations in the dependent and independent variables. Testing different notions of heterogeneity adds that our mechanism seems to be particularly relevant for countries that are politically divided along ethnic lines. Fourth, the dynamics of group participation in the executive are consistent with the assumptions and predictions of the model. Finally, slumps do not seem to be correlated with sudden changes in political institutions, but often occur when weakly institutionalized settings with heterogeneous interest groups are confronted with an external shock.

4 Concluding remarks

This paper presents a political economy theory of declines. It highlights a commitment problem between winners and losers of the recovery process after a crisis, and then analyzes the empirical implications of this theory. We show that it is the combination of ethno-political heterogeneity with weak constraints on the political executive which brings about delayed cooperation. Together, these two factors help to explain why we observe such long declines in some countries and relatively short ones in others.

Both the theory and the empirical analysis suggest that ethnic heterogeneity is harmful for getting groups to agree on a response to a crisis when political institutions are weak. More subtle predictions show that this is mostly an issue of having many powerful groups in the society and does not apply to the same degree when there is a politically dominant group. These findings have important policy implications. On the one hand, political institutions that constrain a country's leadership can contain the adversarial element of ethno-political heterogeneity. On the other hand, our research points out that there can be adverse effects of broad political inclusion when the institutional structure does not sufficiently limit executive power.

While not restricted to understanding declines in Sub-Saharan Africa, we would like to emphasize that we believe these insights are particularly important for understanding the political economy of declines on that subcontinent. Sub-Saharan Africa is home to the countries with the longest and deepest declines, whose politics are often shaped by ethnicity and accompanied by weak institutions governing executive power. While we still need to better understand why ethnic diversity tends to coincide with weak political institutions, we find that there is ample room for managing this heterogeneity better to avoid that welfare gains are lost again in the next crisis.

References

Acemoglu, D. and S. Johnson (2005). Unbundling institutions. *Journal of Political Economy 113*(5), 949–995.

Acemoglu, D. and J. A. Robinson (2006). *Economic Origins of Democracy and Dictatorship*. Cambridge University Press.

Aghion, P., A. Alesina, and F. Trebbi (2004). Endogenous political institutions. *Quarterly Journal of Economics 119*(2), 565–611.

Ahlerup, P. and O. Olsson (2012). The roots of ethnic diversity. *Journal of Economic Growth 17*(2), 71–102.

Alesina, A., R. Baqir, and W. Easterly (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics 114*(4), 1243–1284.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003). Fractionalization. *Journal of Economic Growth 8*(2), 155–194.

Alesina, A. and A. Drazen (1991). Why are stabilizations delayed? *American Economic Review 81*(5), 1170–1188.

Alesina, A. and E. L. Ferrara (2005). Ethnic diversity and economic performance. *Journal of Economic Literature 43*(3), 762–800.

Alesina, A., C. Gennaioli, and S. Lovo (2014, September). Public goods and ethnic diversity: Evidence from deforestation in Indonesia. Working Paper 20504, National Bureau of Economic Research.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2016). Ethnic inequality. *Journal of Political Economy 124*(2), 428–488.

Alesina, A. and E. Zhuravskaya (2011). Segregation and the quality of government in a cross section of countries. *American Economic Review 101*(5), 1872–1911.

Ashraf, Q. and O. Galor (2013). The "Out of Africa" hypothesis, human genetic diversity, and comparative economic development. *American Economic Review 103*(1), 1–46.

Bates, R. H. (2000). Ethnicity and development in Africa: A reappraisal. *American Economic Review 90*(2), 131–134.

Berg, A., J. D. Ostry, and J. Zettelmeyer (2012). What makes growth sustained? *Journal of Development Economics 98*(2), 149–166.

Besley, T. and T. Persson (2011a). Fragile states and development policy. *Journal of the European Economic Association 9*(3), 371–398.

Besley, T. and T. Persson (2011b). The logic of political violence. *Quarterly Journal of Economics 126*(3), 1411–1445.

Bluhm, R., D. de Crombrugghe, and A. Szirmai (2019). Do weak institutions prolong crises? On the identification, characteristics, and duration of declines during economic slumps. *World Bank Economic Review, forthcoming*.

Bueno De Mesquita, B., A. Smith, R. M. Siverson, and J. D. Morrow (2003). *The logic of political survival*. MIT Press.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Padró i Miquel (2015). The value of democracy: Evidence from road building in kenya. *American Economic Review 105*(6), 1817–51.

Caselli, F. and W. J. Coleman (2013). On the theory of ethnic conflict. *Journal of the European Economic Association 11*, 161–192.

Cerra, V. and S. C. Saxena (2008). Growth dynamics: The myth of economic recovery. *American Economic Review 98*(1), 439–57.

Collier, P. (2000). Ethnicity, politics and economic performance. *Economics & Politics 12*(3), 225–245.

Cox, G. W. and M. D. McCubbins (1997). Political structure and economic policy: The institutional determinants of policy outcomes. In S. Haggard and M. D. McCubbins (Eds.), *Presidents, Parliaments, and Policy*, pp. 21–63. Cambridge University Press.

Desmet, K., I. Ortuño-Ortín, and R. Wacziarg (2012). The political economy of linguistic cleavages. *Journal of Development Economics 97*(2), 322–338.

Desmet, K., S. Weber, and I. Ortuño-Ortín (2009). Linguistic diversity and redistribution. *Journal of the European Economic Association 7*(6), 1291–1318.

Drazen, A. and V. Grilli (1993). The benefit of crises for economic reforms. *American Economic Review 83*(3), 598–607.

Easterly, W. (2001). Can institutions resolve ethnic conflict? *Economic Development and Cultural Change 49*(4), 687–706.

Easterly, W., M. Kremer, L. Pritchett, and L. Summers (1993). Good policy or good luck? *Journal of Monetary Economics 32*(3), 459–483.

Easterly, W. and R. Levine (1997). Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics 112*(4), 1203–1250.

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

Esteban, J., L. Mayoral, and D. Ray (2012). Ethnicity and conflict: An empirical study. *American Economic Review 102*(4), 1310–1342.

Esteban, J. and D. Ray (2011). Linking conflict to inequality and polarization. *American Economic Review 101*(4), 1345–1374.

Esteban, J.-M. and D. Ray (1994). On the measurement of polarization. *Econometrica 62*(4), 819–851.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth 8*(2), 195–222.

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

Fernandez, R. and D. Rodrik (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *American Economic Review 81*(5), 1146–1155.

Franck, R. and I. Rainer (2012). Does the leader's ethnicity matter? Ethnic favoritism, education, and health in sub-Saharan Africa. *American Political Science Review 106*(02), 294–325.

Francois, P., I. Rainer, and F. Trebbi (2015). How is power shared in Africa? *Econometrica 83*(2), 465–503.

Gehlbach, S. and E. J. Malesky (2010). The contribution of veto players to economic reform. *Journal of Politics 72*(4), 957–975.

Hausmann, R., L. Pritchett, and D. Rodrik (2005). Growth accelerations. *Journal of Economic Growth 10*(4), 303–329.

Henisz, W. J. (2000). The institutional environment for economic growth. *Economics & Politics 12*(1), 1–31.

Hicken, A., S. Satyanath, and E. Sergenti (2005). Political institutions and economic performance: The effects of accountability and obstacles to policy change. *American Journal of Political Science 49*(4), 897–907.

Hodler, R. and P. A. Raschky (2014). Regional favoritism. *Quarterly Journal of Economics 129*(2), 995–1033.

Jones, B. F. and B. A. Olken (2008). The anatomy of start-stop growth. *Review of Economics and Statistics 90*(3), 582–587.

Labán, R. and F. Sturzenegger (1994). Distributional conflict, financial adaptation and delayed stabilizations. *Economics & Politics 6*(3), 257–276.

Lublin, D. (2017). Electoral systems, ethnic heterogeneity and party system fragmentation. *British Journal of Political Science 47*(2), 373–389.

MacIntyre, A. J. (2003). *The power of institutions: Political architecture and governance.* Cornell University Press.

Michalopoulos, S. (2012). The origins of ethnolinguistic diversity. *American Economic Review 102*(4), 1508–39.

North, D. C., J. J. Wallis, and B. R. Weingast (2009). *Violence and social orders: a conceptual framework for interpreting recorded human history.* Cambridge University Press.

Papell, D. H. and R. Prodan (2014). Long run time series tests of constant steady-state growth. *Economic Modelling 42*, 464 – 474.

Posner, D. N. (2004). Measuring ethnic fractionalization in Africa. *American Journal of Political Science 48*(4), 849–863.

Posner, D. N. (2005). *Institutions and ethnic politics in Africa.* Cambridge University Press.

Rodrik, D. (1999). Where did all the growth go? External shocks, social conflict, and growth collapses. *Journal of Economic Growth 4*(4), 385–412.

Spolaore, E. (2004). Adjustments in different government systems. *Economics &*

Politics 16(2), 117–146.

Svolik, M. W. (2009). Power sharing and leadership dynamics in authoritarian regimes. *American Journal of Political Science 53*(2), 477–494.

Tommasi, M., C. Scartascini, and E. Stein (2014). Veto players and policy adaptability: An intertemporal perspective. *Journal of Theoretical Politics 26*(2), 222–248.

Tsebelis, G. (2002). *Veto players: How political institutions work*. Princeton University Press.

van Reybrouck, D. (2014). *Congo: The epic history of a people*. Harper Collins.

Wimmer, A., L.-E. Cederman, and B. Min (2009). Ethnic politics and armed conflict: A configurational analysis of a new global data set. *American Sociological Review 74*(2), 316–337.

Supplementary Online Appendix to 'Holding on? Ethnic divisions, political institutions and the duration of economic declines' by Richard Bluhm and Kaj Thomsson

A Proofs

Proof of Proposition 1. The utility from cooperation in the first period when the other group cooperates is

$$v_j^1(C,c) = \frac{1}{1-\delta}\left\{(1-2c)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + c(g(0)+g(1))\right\} \tag{A-1}$$

and the utility from choosing to delay cooperation one period when the other group cooperates is

$$v_j^1(D,c) = g((1-\Delta)y_j) + \frac{\delta}{1-\delta}\left\{(1-2p^2)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^2(g(0)+g(1))\right\} \tag{A-2}$$

where $p^2 = c - x$; that is, half the probability of landing outside the safe zone in the second period.

The proof is by contradiction. We conjecture an equilibrium with immediate recovery, such that $v_j^1(C,c) \geq v_j^1(D,c)$. Using $p^2 = c - x$ and rearranging terms, we get

$$\begin{aligned} &g((1-\Delta)y_j) \leq \\ &\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - \left[c + \frac{\delta}{1-\delta}x\right]\left\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - g(0) - g(1)\right\}. \end{aligned} \tag{A-3}$$

Note that concavity implies that $\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - g(0) - g(1)\} > 0$. Inequality (A-3) is contradicted whenever c, x or δ are large enough in relation to Δ, depending on the shape of the utility function $g(y_j)$ and its range, which completes the proof. ■

Proof of Proposition 2. First of all, it is useful to demonstrate that the difference in utility between recovery at any time period (t) and recovery at the subsequent period $(t+1)$ decreases over time. For all $s > t$, we need to check whether

$$v_j^{t+1}(C,c) - v_j^t(C,c) > v_j^{s+1}(C,c) - v_j^s(C,c). \tag{A-4}$$

Note that $v_j^{t+1}(C,c) = v_j^t(D,c)$.

Substituting the utilities and rearranging the inequality, we get

$$\begin{aligned} &g((1-\Delta)y_j) - (1-2p^t)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - p^t(g(0)+g(1)) + \\ &\frac{\delta}{1-\delta}\left\{2(p^t - p^{t+1})\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (p^{t+1} - p^t)(g(0)+g(1))\right\} > \\ &g((1-\Delta)y_j) - (1-2p^s)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - p^s(g(0)+g(1)) + \\ &\frac{\delta}{1-\delta}\left\{2(p^s - p^{s+1})\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (p^{s+1} - p^s)(g(0)+g(1))\right\}. \end{aligned} \tag{A-5}$$

Recall that $p^t = c - (t-1)x$ implies $p^{t+1} - p^t = -x$, so the second and third terms cancel and the inequality reduces to

$$
\begin{aligned}
&(1-2p^t)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^t(g(0)+g(1)) < \\
&(1-2p^s)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^s(g(0)+g(1)).
\end{aligned} \tag{A-6}
$$

Substituting $p^t = c - (t-1)x$ again, it is straightforward to show that this inequality is always satisfied when $s > t$.

Having established this, setting the utility of choosing to cooperate in period t equal to the utility of recovering in period $t+1$ results in an equation that will deliver a potentially non-integer t, such that the smallest higher integer ($\lceil t \rceil$) is the equilibrium time to recovery:

$$
\begin{aligned}
&\frac{1}{1-\delta}\left\{(1-2p^t)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^t(g(0)+g(1))\right\} = g((1-\Delta)y_j) + \\
&\frac{\delta}{1-\delta}\left\{(1-2p^{t+1})\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + p^{t+1}(g(0)+g(1))\right\}.
\end{aligned} \tag{A-7}
$$

Inserting the linear process on $p^t = c - (t-1)x$ yields

$$
\begin{aligned}
&\frac{1}{1-\delta}\left\{(1-2(c-(t-1)x))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-(t-1)x)(g(0)+g(1))\right\} = \\
&g((1-\Delta)y_j) + \frac{\delta}{1-\delta}\left\{(1-2(c-tx))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-tx)(g(0)+g(1))\right\}.
\end{aligned} \tag{A-8}
$$

Isolating the first term of the geometric series gives

$$
\begin{aligned}
&\left\{(1-2(c-(t-1)x))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-(t-1)x)(g(0)+g(1))\right\} + \\
&\frac{\delta}{1-\delta}\left\{(1-2(c-(t-1)x))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-(t-1)x)(g(0)+g(1))\right\} = \\
&g((1-\Delta)y_j) + \frac{\delta}{1-\delta}\left\{(1-2(c-tx))\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + (c-tx)(g(0)+g(1))\right\}
\end{aligned} \tag{A-9}
$$

and after canceling the common terms, we have

$$
\begin{aligned}
&(1-2c)\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] + c(g(0)+g(1)) + \\
&tx\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - (g(0)+g(1))\} \\
&= g((1-\Delta)y_j) + \frac{1}{1-\delta}x\left\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - (g(0)+g(1))\right\}.
\end{aligned} \tag{A-10}
$$

Solving for t^* and simplifying gives

$$
t^* = \frac{g((1-\Delta)y_j) - \mathbb{E}[g(w_j)|w_j \in \mathcal{A}]}{x\{2\mathbb{E}[g(w_j)|w_j \in \mathcal{A}] - (g(0)+g(1))\}} + \frac{c}{x} + \frac{1}{1-\delta}. \tag{A-11}
$$

The proposition follows directly from comparative statics w.r.t. to c

$$\frac{\partial t^*}{\partial c} = \frac{1}{x} > 0 \tag{A-12}$$

that is, stronger executive constraints (smaller c) shorten the time to cooperation. This completes the proof for the interior case. Note that it can also be the case that recovery happens at the point when all uncertainty is resolved, i.e. the point where the probability of being outside the safe zone is zero and no longer changes. If this is the case it is straightforward to see that the time to recovery is shorter with stronger constraints on the executive. This follows directly from the fact, that the time it takes until all uncertainty is resolved is shorter with smaller c. ■

Proof of Proposition 3. Recall that for the asymmetric case only one group risks falling outside the political safe zone. Hence, for there to exist an equilibrium with recovery in the first period, the following condition needs to be true

$$\frac{1}{1-\delta}\left\{(1-p_1(y_j))y_j + p_1(y_j)z\right\} \geq (1-\Delta)y_j + \frac{\delta}{1-\delta}y_j \tag{A-13}$$

which simplifies to

$$\Delta y_j + \frac{1}{1-\delta}\left\{p_1(y_j)(z-y_j)\right\} \geq 0. \tag{A-14}$$

An decrease in concentration (asymmetry) makes delay more likely if the left hand side of the inequality is a decreasing function of y_j. This is true when the derivative of the left hand side is negative:

$$\Delta + \frac{1}{1-\delta}\left\{\frac{dp_1(y_j)}{dy_j}(z-y_j) - p_1(y_j)\right\} < 0 \tag{A-15}$$

which completes the proof. ■

Proof of Proposition 4. Remember that there always exists an equilibrium with recovery in period two in the sub-game that starts in period two after delay in period one. If all of the other groups decide to cooperate in period 1, it is optimal for the remaining group to cooperate if the following condition holds

$$\begin{aligned}&\frac{1}{1-\delta}\left\{(1-Jp_1(y_j))\frac{1}{J} + (J-1)p_1(y_j)\left(\frac{1}{J}+\frac{1}{(J-1)J}\right) + p_1(y_j)z\right\} \geq \\ &(1-\Delta)\frac{1}{J} + \frac{\delta}{1-\delta}\frac{1}{J}.\end{aligned} \tag{A-16}$$

The second term inside the curly braces simplifies to $p_t(y_j)$, so that inequality (A-16) becomes

$$\frac{1}{1-\delta}\left\{(1-Jp_1(y_j))\frac{1}{J}+p_1(y_j)+p_1(y_j)z\right\}\geq(1-\Delta)\frac{1}{J}+\frac{\delta}{1-\delta}\frac{1}{J} \tag{A-17}$$

or

$$\frac{1}{1-\delta}\left\{\frac{1}{J}+p_1(y_j)z\right\}\geq(1-\Delta)\frac{1}{J}+\frac{\delta}{1-\delta}\frac{1}{J} \tag{A-18}$$

and, after some algebraic manipulation, this simplifies to

$$\frac{\Delta}{J}+\frac{1}{1-\delta}p_1(y_j)z\geq 0. \tag{A-19}$$

Now remember that $p_1(y_j)$ is increasing in J, as symmetry implies $y_j = 1/J$ and $z < 0$. As a result, the inequality becomes harder to satisfy if the number of groups increases, proving part *i)* of Proposition 4.

For part *ii)*, consider first a setting with $c = 0$. Then $p_1(y_j) = 0$ for all j and delay will not occur. Now introduce $c > 0$; in this case $p_1(y_j) \geq p_1(y_i)$ for all $j > i$ by assumption, and from part *i)* of this proposition we know that an increase in the number of groups makes delay more likely. Hence, the introduction of a threshold c greater than zero is more likely to imply delay when the number of groups is larger. ■

B Identifying slumps

Bluhm et al. (2019) outline a new approach to finding the duration of the decline phase of large economic slumps. The restricted structural change approach is a variant of Bai (1997) and Papell and Prodan (2014). Bluhm et al. (2019) provide a formal description of the break search algorithm and bootstrap. The companion paper also reports a variety of perturbations for each of the key parameters used in the break search. Here we only summarize the approach.

We specify the following model for each (log) GDP per capita time series y_t:

$$y_t = \alpha + \beta t + \gamma_0 \mathbf{1}(t > t_{b1}) + \gamma_1 (t - t_{b1}) \mathbf{1}(t > t_{b1}) + \gamma_2 (t - t_{b2}) \mathbf{1}(t > t_{b2}) + \sum_{i=1}^{p} \delta_i y_{t-i} + \epsilon_t$$

where t_{b1} and t_{b2} are the endogenous break dates, $\mathbf{1}(\cdot)$ is an indicator function, and p is the lag order. The optimal $AR(p)$ model is determined by the Bayesian information criterion (BIC). We require that $t_{b2} \geq t_{b1} + 4$, so that the period between two successive breaks is at minimum 4 years.

We impose two restrictions. First, we require $\beta > 0$, so that growth must be positive in the years before a slump begins. Second, we also impose the condition that $\gamma_0 < 0$, so that a slump always starts with a drop in the intercept. Slope shifts are left unrestricted, so that the model can catch unfinished slumps (e.g., declines from t_{b1} onwards, possibly lasting until the end of a country's time series). Next we compute the the sup-W test statistic of the null of no break versus two breaks ($\mathbb{H}_0 : \gamma_0 = \gamma_1 = \gamma_2 = 0$).

Note that the individual Wald tests over which the sup-W statistic is computed are not statistically independent. Hence, we bootstrap the empirical p-value of the sup-W statistic using a recursive bootstrap (Diebold and Chen, 1996). If the bootstrap test rejects at the desired significance level, α, we record the break pair $(\widehat{t}_{b1}, \widehat{t}_{b2})$ and split the sample into a series running until the first break and a series starting just after the second break. The process starts again on each sub-sample until the bootstrap test fails to reject the null hypothesis of no breaks or the sample gets too small ($T < 20$). We fix the nominal level at $\alpha = 0.1$ for the baseline results.

A slump is completed when the level of GDP per capita has at least caught up again with its own past. If that point is reached within the sample, we define the recovery to have been completed in the first year $t_c > \widehat{t}_{b1}$ where $y_{t_c} \geq y_{\widehat{t}_{b1}}$.

Next we date the though. The pre-slump level of GDP per capita is not always reached again within the sample period. In that case, the duration of the slump is censored. Even though GDP per capita may be recovering, we do not know how long it will take to restore the earlier peak. A provisional trough is then observed when y_t attains a minimum after

$\widehat{t}_{b1}$. To cover all cases, we estimate the trough to have occurred at time:

$$\widehat{t}_{min} \quad = \quad \begin{cases} \operatorname{argmin}_{j\in(\widehat{t}_{b1},t_c]} y_j & \text{if the spell is completed in year } t_c, \\ \operatorname{argmin}_{j\in(\widehat{t}_{b1},T]} y_j & \text{if the spell is censored.} \end{cases} \tag{B-1}$$

Last but not least, we define the duration of the decline phase as the duration of the beginning of the slump until the (provisional) trough, or $\widetilde{t}_D = \widehat{t}_{min} - \widehat{t}_{b1}$.

Table B-1 summarizes the breaks found when running this algorithm on all countries with at least 20 data points and more than one million inhabitants in the Penn World Table 7.0 (series *rgdpch*).

Table B-1: Estimated breaks with troughs: 58 Episodes

Code	T_0	$\widehat{t}_{b1}$	$\widehat{t}_{min}$	$\widehat{t}_{b2}$	T	Sup-W	Critical W	p-value	Drop (%)	Duration
ALB	1970	1990	1991	2002	2008	18.5	13.6	0.007	-15.32	1
ARE	1986	1990	1999	2002	2008	29.1	14.5	0.003	-10.90	9
AUS	1950	1954	1957	1966	2008	8.3	8.7	0.064	-0.72	3
AUS	1967	1989	1991	1998	2008	10.1	10.7	0.059	-2.29	2
BDI	1960	1971	1972	1988	2008	9.9	11.3	0.089	-3.23	1
BEL	1950	1957	1958	1973	2008	12.8	12.1	0.029	-2.24	1
BGR	1970	1988	1997	1997	2008	16.3	12.8	0.010	-23.79	9
BHR	1970	1980	1987	1986	2008	14.4	11.0	0.010	-44.12	7
BRA	1950	1980	1983	2003	2008	12.5	12.3	0.043	-14.60	3
CAF	1960	1978	2005	2005	2008	8.3	8.7	0.060	-46.38	27
CHE	1950	1974	1975	1978	2008	10.7	10.6	0.047	-7.87	1
CHL	1951	1953	1954	1972	1973	12.0	8.5	0.017	-9.06	1
CHL	1951	1974	1975	1979	1980	13.3	10.8	0.021	-16.50	1
CHL	1951	1981	1983	1995	2008	12.6	11.4	0.025	-21.22	2
CHN	1952	1960	1962	1977	2008	13.9	12.9	0.029	-23.71	2
CMR	1960	1986	1995	1990	2008	12.0	12.3	0.055	-40.46	9
COG	1960	1974	1977	1982	2008	11.9	12.5	0.069	-21.35	3
CRI	1950	1955	1956	1963	1979	11.4	11.3	0.048	-4.39	1
CRI	1950	1980	1982	2002	2008	17.2	10.6	0.002	-17.47	2
CUB	1970	1988	1993	1995	2008	11.4	12.5	0.072	-34.70	5
CYP	1950	1973	1975	1977	2008	15.5	9.7	0.001	-31.40	2
CYP	1978	1990	1991	1995	2008	11.6	14.6	0.098	-10.19	1
DNK	1950	1954	1955	1965	2008	12.9	11.7	0.022	-1.56	1
DZA	1960	1984	1994	1996	2008	10.9	8.2	0.013	-14.09	10
ETH	1950	1972	1992	1993	2008	11.5	10.2	0.020	-30.68	20
FIN	1950	1989	1993	2006	2008	10.6	10.8	0.057	-16.34	4
GAB	1960	1976	1987	1997	2008	10.6	11.2	0.062	-50.56	11
GMB	1960	1982	1998	2002	2008	16.4	11.2	0.006	-25.33	16
GRC	1951	1973	1974	1994	2008	17.9	11.6	0.003	-6.92	1
GTM	1950	1980	1988	1984	2008	15.1	12.3	0.015	-19.14	8
HUN	1970	1990	1992	2004	2008	15.6	13.5	0.018	-10.56	2
IDN	1960	1997	1999	2001	2008	13.5	10.6	0.013	-17.49	2
IRN	1955	1976	1981	1980	2008	15.9	11.6	0.004	-56.78	5
IRQ	1970	1990	2003	1994	2008	9.1	8.9	0.046	-66.43	13
JPN	1950	1973	1974	1990	2008	13.5	13.4	0.050	-2.85	1
MEX	1950	1981	1988	1995	2008	11.9	11.0	0.038	-17.03	7
MNG	1970	1990	1993	2003	2008	46.5	11.7	0.000	-41.81	3
MOZ	1960	1981	1986	1995	2008	12.6	12.0	0.037	-24.99	5

Continued on next page

Table B-1 – *Continued from previous page*

Code	T_0	$\widehat{t}_{b1}$	$\widehat{t}_{min}$	$\widehat{t}_{b2}$	T	Sup-W	Critical W	p-value	Drop (%)	Duration
MYS	1955	1984	1986	1993	2008	9.1	10.5	0.093	-7.47	2
NPL	1960	1979	1980	2000	2008	10.6	8.9	0.025	-5.33	1
NZL	1950	1974	1978	1992	2008	9.9	10.5	0.070	-9.03	4
OMN	1970	1979	1980	1985	2008	12.4	9.0	0.007	-21.61	1
PER	1950	1958	1959	1966	1976	11.9	9.3	0.022	-6.91	1
PER	1950	1977	1992	1992	2008	11.0	10.3	0.037	-29.30	15
PHL	1950	1983	1985	2003	2008	12.8	10.2	0.007	-16.78	2
POL	1970	1979	1982	1993	2008	13.8	12.1	0.027	-22.55	3
PRY	1980	1989	2002	2002	2008	8.8	8.8	0.049	-14.24	13
RWA	1960	1993	1994	1997	2008	18.0	7.9	0.001	-45.38	1
SAU	1986	1992	1999	2002	2008	14.6	13.3	0.039	-18.75	7
SLE	1961	1995	1999	2006	2008	14.2	11.1	0.011	-41.65	4
SLV	1950	1978	1983	1987	2008	18.2	10.2	0.002	-25.82	5
TGO	1960	1979	2008	1989	2008	9.6	10.1	0.065	-53.60	29
THA	1950	1996	1998	2003	2008	10.7	7.8	0.003	-14.17	2
TTO	1950	1961	1963	1969	1981	16.8	14.9	0.020	-0.78	2
TTO	1950	1982	1993	2006	2008	12.4	12.6	0.054	-28.96	11
UGA	1950	1977	1986	1987	2008	11.6	10.5	0.029	-30.27	9
USA	1950	1957	1958	1966	2008	8.7	9.3	0.075	-2.51	1
ZMB	1955	1968	2001	2000	2008	15.0	10.9	0.007	-68.99	33

Notes(s): Out of a total of 70 episodes identified by the sequential algorithm, 12 are not valid slumps. We discard invalid episodes that are driven by positive breaks in the slope coefficient but fail the negative growth criterion due to the presence of the auto-regressive terms. A simple rule is applied to these cases, requiring that an actual contraction occurs within the range of the two estimated breaks, otherwise there is no slump. The invalid episodes are [country code (spell number)]: AUT (1), AUT (2), CHN (1), FIN (1), HKG (1), IRN (1), MRT (1), PRY (1), TZA (1).

C Additional summary statistics

This section provides additional summary statistics for the data used in the main text and in this Online Appendix.

Table C-1 shows summary statistics of the independent variables used in the main specifications.

Table C-2 adds the pairwise correlations among these variables, including the duration of the decline phase, with and without logs. Note that the partial correlations of the duration and our main explanatory variables emphasized in the paper account for the correlation between duration and GDP per capita, among other correlates.

Table C-4 mirrors the summary statistics of the dependent variable from the main text but shows the results for a larger set of slumps estimated using a more lenient nominal size threshold of 20 percent. As before, we obtain a balanced distribution of slumps across continents. More or less half of all countries in each region experience a slump, suggesting that geographic selection is limited.

Table C-4 adds the summary statistics of the independent variables for the larger sample used in Table D-1 in this Online Appendix.

Table C-1: Summary statistics of right hand side variables

VARIABLES	Obs	Mean	Std. dev.	Min	Max
$XCONST_0$	58	3.48	2.49	1.00	7.00
ELF	58	45.39	33.71	0.07	95.98
ELF_0	57	36.00	25.71	0.00	80.39
POL	58	40.04	24.98	0.14	85.99
POL_0	57	19.35	16.42	0.00	56.95
ELA_0	57	48.75	33.40	0.10	100.00
$GROUPS_0$	57	4.19	6.43	0.00	47.00
$EGIPGRPS_0$	57	1.37	1.33	0.00	7.00
$EXCLGRPS_0$	57	2.33	6.17	0.00	46.00
$MONPOP_0$	57	21.48	35.94	0.00	97.30
$DOMPOP_0$	57	21.28	34.25	0.00	98.00
$\ln GDPPC_0$	58	8.53	1.21	5.87	10.63

Note(s): The table shows summary statistics of the independent variables used in the main text.

Table C-2: Correlation matrix

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) $XCONST_0$	1.00													
(2) ELF	-0.26	1.00												
(3) ELF_0	-0.35	0.67	1.00											
(4) POL	0.10	0.58	0.36	1.00										
(5) POL_0	0.01	-0.42	-0.55	-0.20	1.00									
(6) ELA_0	0.36	-0.62	-0.93	-0.35	0.31	1.00								
(7) $GROUPS_0$	-0.13	0.19	0.38	0.15	-0.38	-0.20	1.00							
(8) $EGIPGRPS_0$	-0.00	0.37	0.47	0.09	-0.32	-0.39	0.12	1.00						
(9) $EXCLGRPS_0$	-0.15	0.08	0.24	0.13	-0.26	-0.11	0.96	-0.07	1.00					
(10) $MONPOP_0$	-0.03	-0.36	-0.29	-0.05	0.22	0.29	0.07	-0.17	0.12	1.00				
(11) $DOMPOP_0$	0.20	-0.23	-0.27	-0.23	0.10	0.27	-0.04	-0.17	-0.04	-0.38	1.00			
(12) $\ln GDPPC_0$	0.43	-0.28	-0.41	0.13	0.13	0.35	-0.34	-0.07	-0.34	0.04	0.02	1.00		
(13) $\tilde{t}$	-0.35	0.46	0.36	0.09	-0.08	-0.41	-0.01	0.41	-0.07	-0.18	-0.14	-0.24	1.00	
(14) $\ln \tilde{t}$	-0.39	0.47	0.42	0.14	-0.07	-0.49	0.00	0.34	-0.05	-0.25	-0.11	-0.13	0.88	1.00

Note(s): The table shows a matrix of pairwise correlations among the variables used in the main text.

Table C-3: Summary statistics of slumps, larger sample

	Africa	Americas	Asia	Europe	Oceania	World
Countries	46	25	35	29	3	138
Countries with slumps	22	14	19	13	2	70
Number of slumps	30	19	20	16	4	89
Total years in decline	258	111	73	37	10	489
Duration of decline:						
– Min	1	1	1	1	1	1
– Median	5	3	2	2	2	3
– Mean	8.60	5.84	3.65	2.31	2.50	5.49
– Max	33	24	13	9	4	33
Incidence Rate	0.09	0.15	0.23	0.43	0.40	0.16

Note(s): The table shows summary statistics of the duration of the decline phase of slumps based on the larger sample. The larger sample of slumps was obtained by running the structural break algorithm with a more lenient nominal size threshold of 20%. A few countries (e.g. Chile) have repeated slumps which generates the discrepancy between the reported number of countries and number slumps. The incidence rate is defined as the number of exits from the decline period over the total years in decline.

Table C-4: Summary statistics of right hand side variables, larger sample

Variables	Obs	Mean	Std. dev.	Min	Max
$XCONST_0$	83	3.61	2.52	1.00	7.00
ELF	89	48.45	34.00	0.07	95.98
ELF_0	82	37.28	27.74	0.00	86.49
POL	89	39.83	23.96	0.14	85.99
POL_0	82	16.08	16.24	0.00	56.95
ELA_0	82	49.10	34.53	0.10	100.00
$GROUPS_0$	82	4.27	5.81	0.00	47.00
$EGIPGRPS_0$	82	1.61	1.94	0.00	14.00
$EXCLGRPS_0$	82	2.20	5.26	0.00	46.00
$MONPOP_0$	82	18.22	34.07	0.00	97.30
$DOMPOP_0$	82	19.38	33.70	0.00	98.00
$\ln GDPPC_0$	89	8.31	1.23	5.87	10.63

Note(s): The table shows summary statistics of the independent variables used in the main text based on the larger sample. The larger sample of slumps was obtained by running the structural break algorithm with a more lenient nominal size threshold of 20%.

D Additional survival regressions

In this section, we vary all important decisions underlying the tables presented in the main text and further illustrate the robustness of the empirical results.

Left hand side variations. Table D-1 expands the sample size by raising the type II error threshold used during the break search. This increases the sample size at the expense of detecting more false positives (smaller recessions, not slumps). The results are qualitatively unchanged and even quantitatively very similar.

Table D-2 changes the data used during the break search to the version 9 vintage of the Penn World Tables running from 1950 to 2014 (series *rgdpna/pop*). Note that the computation of the headline GDP per capita series was changed substantially over these vintages, so that the number of slumps and country coverage is not the same across these data sets. In spite of these qualifications, the results mirror our main findings.

Table D-3 changes the resampling technique from a recursive parametric bootstrap to Hansen's (2000) fixed-design bootstrap which allows for nonstationarity, lagged dependent variables, and conditional heteroskedasticity. Table D-4 uses the same bootstrap with a larger nominal size of 20%. Both of these changes result in substantially larger samples (with up to 106 spells) while our core results remain intact.

Right hand side variations. Table D-5 alters the measure of ethnic fractionalization used in the main specification to other standard data sources used in the literature, that is, Alesina et al. (2003), Fearon (2003) and the *Atlas Narodov Mira*. The results are very similar across the different measures, with the exception of religious fragmentation.

Table D-6 expands on this theme and includes measures of diversity, as well as measures capturing different dimensions of segregation. A similar pattern emerges. Only genetic diversity and religious segregation do not generate an significant interaction term.

Table D-7 uses measures of spatial and ethnic inequality instead of fractionalization. Here the results are not unambiguous, although they tend to point in the same direction. Taken together, these three tables underline that our findings are not specific to ethnic diversity but instead capture the existence of identity groups more broadly.

Table D-8 alters the data used to measure the political constraints placed on the executive. Executive constraints are now measured using the data provided by Henisz (2000). The results remain robust to this change, especially when the variable *POLCON III* is considered. We are nonetheless reluctant to emphasize these results, as the data are derived from a veto player model with a very different focus than our theory.

Table D-9 presents two further set of perturbations. The first columns vary the structure of the time effects to initial decade dummies (with a different origin), seven year dummies, and quinquennial dummies. The last three columns include several measures of heterogeneity and their interactions at the same time. The results are robust to these alterations, not counting the strong collinearity introduced in the second to last column.

Table D-10 alters the functional form assumed by the event history model and shows that our findings are not specific to the lognormal AFT model.

Last but not least, Table D-11 adds a set of policy variables to investigate whether these decisions affect the recovery likelihood. Government size, changes in the exchange rate regime and programs by the two largest international financial institutions have the expected sign and there is some indication that they accelerate the recovery process. Our main results remain unaltered when these variables are added, suggesting that our mechanism operates more broadly through a variety of policies, their implications, and even policy inaction.

Table D-1: Robustness – Sample of slumps: larger sample

	Dependent variable: $\ln \tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Ethnologue			*Ethnic Power Relations*		
$XCONST_0$	-0.195***	-0.245***	-0.187***	-0.180***	-0.220***	-0.173***
	(0.050)	(0.054)	(0.051)	(0.054)	(0.051)	(0.053)
ELF	0.013***	0.014***	0.017***			
	(0.004)	(0.003)	(0.004)			
$XCONST_0 \times ELF$		-0.003**				
		(0.001)				
POL			-0.007			
			(0.006)			
ELF_0				0.013**	0.014***	0.015***
				(0.005)	(0.005)	(0.005)
$XCONST_0 \times ELF_0$					-0.003**	
					(0.001)	
POL_0						0.008
						(0.008)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Decade dummies	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	71	71	71	69	69	69
Spells	83	83	83	81	81	81
Years of decline	468	468	468	464	464	464
Log-$\mathcal{L}$	-114.133	-111.929	-113.498	-114.993	-113.526	-114.556
Pseudo-R^2	0.093	0.110	0.098	0.068	0.080	0.072

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The larger sample of slumps was obtained by running the structural break algorithm with a more lenient nominal size threshold of 20%. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-2: Robustness – Penn World Table 9 until 2014

	Dependent variable: $\ln \tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Ethnologue			*Ethnic Power Relations*		
$XCONST_0$	-0.196***	-0.264***	-0.178***	-0.171**	-0.216***	-0.162**
	(0.067)	(0.076)	(0.067)	(0.075)	(0.076)	(0.073)
ELF	0.015***	0.017***	0.021***			
	(0.004)	(0.005)	(0.007)			
$XCONST_0 \times ELF$		-0.004**				
		(0.002)				
POL			-0.015*			
			(0.008)			
ELF_0				0.013**	0.016***	0.020***
				(0.006)	(0.006)	(0.007)
$XCONST_0 \times ELF_0$					-0.004**	
					(0.002)	
POL_0						0.016
						(0.011)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	43	43	43	42	42	42
Spells	54	54	54	52	52	52
Years of decline	383	383	383	367	367	367
Log-$\mathcal{L}$	-77.010	-75.820	-75.396	-76.545	-75.613	-75.826
Pseudo-R^2	0.089	0.103	0.108	0.054	0.066	0.063

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The sample of slumps was obtained by running the structural break algorithm on the series *rgdpna/pop* from the Penn World Table 9.0. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-3: Robustness – Hansen's fixed design bootstrap

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: $\ln \tilde{t}$					
	Ethnologue			*Ethnic Power Relations*		
$XCONST_0$	-0.151***	-0.184***	-0.147***	-0.146***	-0.217***	-0.145***
	(0.055)	(0.068)	(0.055)	(0.048)	(0.070)	(0.048)
ELF	0.013***	0.015***	0.017***			
	(0.004)	(0.004)	(0.005)			
$XCONST_0 \times ELF$		-0.002*				
		(0.001)				
POL			-0.008			
			(0.006)			
ELF_0				0.027***	0.027***	0.027***
				(0.005)	(0.004)	(0.005)
$XCONST_0 \times ELF_0$					-0.004**	
					(0.002)	
POL_0						0.001
						(0.006)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	66	66	66	62	62	62
Spells	76	76	76	72	72	72
Years of decline	389	389	389	370	370	370
Log-$\mathcal{L}$	-103.499	-101.981	-102.713	-90.595	-87.741	-90.584
Pseudo-R^2	0.087	0.100	0.094	0.158	0.185	0.158

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The larger sample of slumps was obtained by running the structural break algorithm with Hansen's fixed design bootstrap (Hansen, 2000). The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-4: Robustness – Hansen's fixed design bootstrap, larger sample

	Dependent variable: $\ln \tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Ethnologue			*Ethnic Power Relations*		
$XCONST_0$	-0.147***	-0.164***	-0.144***	-0.150***	-0.188***	-0.152***
	(0.051)	(0.054)	(0.052)	(0.047)	(0.047)	(0.048)
ELF	0.010***	0.010***	0.012***			
	(0.004)	(0.004)	(0.004)			
$XCONST_0 \times ELF$		-0.002				
		(0.001)				
POL			-0.004			
			(0.005)			
ELF_0				0.016***	0.017***	0.015***
				(0.004)	(0.004)	(0.005)
$XCONST_0 \times ELF_0$					-0.003**	
					(0.001)	
POL_0						-0.002
						(0.007)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	93	93	93	88	88	88
Spells	106	106	106	101	101	101
Years of decline	596	596	596	573	573	573
Log-$\mathcal{L}$	-151.047	-150.127	-150.730	-142.153	-139.801	-142.118
Pseudo-R^2	0.051	0.057	0.053	0.067	0.082	0.067

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The larger sample of slumps was obtained by running the structural break algorithm with Hansen's fixed design bootstrap (Hansen, 2000) and by employing a more lenient nominal size threshold of 20%. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-5: Robustness – Measures of fractionalization

	Dependent variable: $\ln \tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Alesina et al.			*Fearon*		*Atlas*
$XCONST_0$	-0.230***	-0.293***	-0.184**	-0.239***	-0.243***	-0.284***
	(0.069)	(0.082)	(0.073)	(0.073)	(0.053)	(0.073)
Ethnic (H^A)	0.020***					
	(0.006)					
$XCONST_0 \times H^A$	-0.004**					
	(0.002)					
Linguistic (H^B)		0.021***				
		(0.006)				
$XCONST_0 \times H^B$		-0.004***				
		(0.002)				
Religious (H^C)			0.005			
			(0.008)			
$XCONST_0 \times H^C$			-0.004*			
			(0.002)			
Ethnic (H^D)				0.019***		
				(0.006)		
$XCONST_0 \times H^D$				-0.005***		
				(0.002)		
Cultural (H^E)					0.028***	
					(0.005)	
$XCONST_0 \times H^E$					-0.008***	
					(0.002)	
Ethnic (H^F)						0.020***
						(0.005)
$XCONST_0 \times H^F$						-0.005***
						(0.002)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Decade dummies	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	48	45	48	48	48	45
Spells	58	55	58	58	58	55
Years of decline	348	337	348	348	348	333
Log-$\mathcal{L}$	-63.681	-55.225	-67.932	-63.073	-58.602	-58.670
Pseudo-R^2	0.275	0.341	0.227	0.282	0.333	0.298

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The measures of heterogeneity H^A to H^C are obtained from Alesina et al. (2003), H^D and H^E are from Fearon (2003), and H^F is from the Atlas Narodov Mira as published by Alesina et al. (2003). The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-6: Robustness – Alternate measures of heterogeneity

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: $\ln \tilde{t}$					
	Desmet et al.		*Ashraf & Galor*	*Alesina & Zhuravskaya*		
$XCONST_0$	-0.178***	-0.164***	-0.144**	-0.235**	-0.337***	-0.066
	(0.049)	(0.049)	(0.065)	(0.111)	(0.102)	(0.167)
Greenberg (H^A)	0.027***					
	(0.008)					
$XCONST_0 \times H^A$	-0.008***					
	(0.003)					
Peripheral Het. (H^B)		0.033**				
		(0.013)				
$XCONST_0 \times H^B$		-0.011***				
		(0.004)				
Genetic Div. (H^C)			0.065			
			(0.066)			
$XCONST_0 \times H^C$			-0.002			
			(0.009)			
Ethnic Seg. (H^D)				0.033		
				(0.021)		
$XCONST_0 \times H^D$				-0.011		
				(0.009)		
Linguistic Seg. (H^E)					0.032**	
					(0.015)	
$XCONST_0 \times H^E$					-0.014**	
					(0.007)	
Religious Seg. (H^F)						0.049**
						(0.024)
$XCONST_0 \times H^F$						0.019
						(0.031)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Decade dummies	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	48	48	48	31	32	25
Spells	58	58	58	38	38	32
Years of decline	348	348	348	249	248	235
Log-$\mathcal{L}$	-62.468	-63.393	-68.777	-39.815	-38.150	-34.907
Pseudo-R^2	0.289	0.278	0.217	0.342	0.366	0.307

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The measures of linguistic heterogeneity H^A and H^B are obtained from Desmet et al. (2009), genetic diversity H^C is from Ashraf and Galor (2013), and the segregation measures H^D to H^F are from Alesina and Zhuravskaya (2011). The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-7: Robustness – Measures of spatial and ethnic inequality

	Dependent variable: $\ln \tilde{t}$					
Variables	(1)	(2)	(3)	(4)	(5)	(6)
$XCONST_0$	-0.148**	-0.158**	-0.157**	-0.186***	-0.181**	-0.165**
	(0.072)	(0.065)	(0.063)	(0.070)	(0.073)	(0.077)
Spatial (H^A)	0.005					
	(0.008)					
$XCONST_0 \times H^A$	-0.001					
	(0.002)					
Ethnic Level 1 (H^B)		0.012**				
		(0.005)				
$XCONST_0 \times H^B$		-0.004*				
		(0.002)				
Ethnic Level 5 (H^C)			0.007			
			(0.005)			
$XCONST_0 \times H^C$			-0.002			
			(0.001)			
Ethnic Level 10 (H^D)				0.008*		
				(0.005)		
$XCONST_0 \times H^D$				-0.003*		
				(0.001)		
Ethnic Level 15 (H^E)					0.007	
					(0.005)	
$XCONST_0 \times H^E$					-0.002	
					(0.001)	
Ethnic GREG (H^F)						0.014*
						(0.007)
$XCONST_0 \times H^F$						-0.002
						(0.002)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Decade dummies	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	48	48	48	48	48	48
Spells	58	58	58	58	58	58
Years of decline	348	348	348	348	348	348
Log-$\mathcal{L}$	-68.875	-65.912	-67.546	-66.854	-67.546	-66.702
Pseudo-R^2	0.216	0.250	0.231	0.239	0.231	0.241

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The measures of spatial and ethnic inequality H^A to H^F are obtained from Alesina et al. (2016). The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-8: Robustness – Measures of political constraints

	Dependent variable: $\ln \tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Henisz Political Constraints					
ELF	0.017***	0.020***	0.021***	0.021***	0.022***	0.022***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
POLCON III	-1.317**	-2.130***				
	(0.568)	(0.631)				
POLCON III × *ELF*		-0.041***				
		(0.016)				
POLCON V			-0.901**	-1.092**		
			(0.399)	(0.477)		
POLCON V × *ELF*				-0.009		
				(0.010)		
POLCON VJ					-1.076	-2.289**
					(0.730)	(1.019)
POLCON VJ × *ELF*						-0.027*
						(0.016)
	Control sets					
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Decade dummies	Yes	Yes	Yes	Yes	Yes	Yes
	Summary stats					
Exits	47	47	39	39	34	34
Spells	57	57	49	49	44	44
Years of decline	347	347	335	335	325	325
Log-$\mathcal{L}$	-62.983	-60.722	-50.894	-50.602	-45.363	-44.934
Pseudo-R^2	0.269	0.295	0.304	0.308	0.305	0.311

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. Executive constraints are measured using the data from Henisz (2000). *POLCONIII* is derived from a structural veto-player model. *POLCONV* adds two additional veto points for the judiciary and sub-federal entities. *POLCONVJ* includes measures of alignment and fractionalization of the High Court. Only *POLCONIII* is still remotely related to the parameter c in our model. However, these measures always include legislative fractionalization, while we are concerned with ethnic fractionalizaton of the executive. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-9: Robustness – Time effects and additional interactions

	Dependent variable: $\ln \tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Varying the time effects			*More heterogeneity interactions*		
$XCONST_0$	-0.181***	-0.225***	-0.160***	-0.254***	-0.267***	-0.238***
	(0.064)	(0.065)	(0.059)	(0.087)	(0.067)	(0.071)
ELF	0.021***	0.023***	0.019***	0.019***	0.016*	0.017**
	(0.004)	(0.004)	(0.004)	(0.003)	(0.009)	(0.009)
$XCONST_0 \times ELF$	-0.003***	-0.004***	-0.003***	-0.004*	-0.005	-0.004***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.001)
$XCONST_0 \times POL$				0.000		
				(0.002)		
$XCONST_0 \times ELF$ (Fearon)					0.000	
					(0.006)	
POL						0.006
						(0.009)
ELF (Fearon)					0.005	0.007
					(0.010)	(0.009)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Time dummies	10 yrs.*	7 yrs.	5 yrs.	10 yrs.	10 yrs.	10 yrs.
Summary stats						
Exits	48	48	48	48	45	44
Spells	58	58	58	58	55	54
Years of decline	348	348	348	348	333	331
Log-$\mathcal{L}$	-58.565	-56.115	-53.624	-58.133	-55.623	-53.101
Pseudo-R^2	0.333	0.361	0.390	0.338	0.335	0.355

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. Columns (1) to (3) vary the origin (marked *) and sequence of the time effects. Columns (3) to (6) simultaneously include other measures of heterogeneity as interactions or in levels. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-10: Robustness – Functional form

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: $\ln \tilde{t}$					
	Coefficients ($\mathbb{H}_0 = 0$)		*Hazard ratios* ($\mathbb{H}_0 = 1$)			
	Log-logistic		*Weibull*		*Cox*	
$XCONST_0$	-0.270***	-0.253***	1.455***	1.547***	1.363***	1.412***
	(0.084)	(0.075)	(0.126)	(0.147)	(0.113)	(0.115)
ELF	0.020***	0.020***	0.968***	0.963***	0.973***	0.971***
	(0.004)	(0.004)	(0.008)	(0.009)	(0.007)	(0.008)
$XCONST_0 \times ELF$	-0.004**	-0.004***	1.004*	1.008***	1.004*	1.006***
	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Decade dummies	No	Yes	No	Yes	No	Yes
Summary stats						
Exits	48	48	48	48	48	48
Spells	58	58	58	58	58	58
Years of decline	348	348	348	348	348	348
Log-$\mathcal{L}$	-64.578	-59.569	-66.091	-59.615	-140.459	-134.788
Pseudo-R^2	0.274	0.330	0.306	0.374	0.127	0.163

Note(s): The table shows the results from survival regressions of the duration of economic declines on our variables of interest. Columns (1) and (2) are AFT models with a log-logistic density, columns (3) and (4) are hazard models with a Weibull hazard, and columns (5) and (6) are non-parametric proportional hazard models. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D-11: Robustness – Policy variables

	Dependent variable: $\ln\tilde{t}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Only policy variables			*With preferred specification*		
Government share of GDP_t	-0.031			0.007		
	(0.024)			(0.019)		
Change in exchange rate regime_t		-0.670**			-0.578**	
		(0.318)			(0.250)	
No. of IMF projects_{t-1}			0.268			0.083
			(0.291)			(0.236)
No. of World Bank projects_{t-1}			-0.097**			-0.083*
			(0.049)			(0.044)
$XCONST_0$				-0.252***	-0.251***	-0.185***
				(0.060)	(0.059)	(0.072)
ELF				0.020***	0.019***	0.023***
				(0.004)	(0.003)	(0.004)
$XCONST_0 \times ELF_0$				-0.004***	-0.004***	-0.003
				(0.001)	(0.001)	(0.002)
Control sets						
GDP per capita	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Decade dummies	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Exits	48	47	32	48	47	32
Spells	58	57	42	58	57	42
Years of decline	348	343	312	348	343	312
Log-$\mathcal{L}$	-70.926	-69.149	-50.536	-58.079	-55.345	-40.153
Pseudo-R^2	0.193	0.203	0.184	0.339	0.362	0.352

Note(s): The table shows the results from log-normal survival regressions of the duration of economic declines on our variables of interest. The government share of GDP is taken from the Penn World Table 7.0. A change in the exchange rate regime is coded as unity when there is a jump of any size in the coarse regime classification provided by Ilzetzki et al. (2017). Ilzetzki et al. (2017) code regimes in five categories from 'fixed' over 'managed float' to 'freely falling'. Data on the number of IMF and World Bank programs are from Boockmann and Dreher (2003) and Dreher (2006). We count all IMF programs recorded in the data and all WB programs, except environmental projects. The data are lagged by one year to allow some time for projects to take effect. The standard errors are clustered on the country level to account for repeated spells. All specifications include a constant (not shown). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E Additional group level regressions

Table E-1 revisits a core assumption of the theory presented in the paper, that is, groups face an elevated probability of falling out of government during a crisis, particularly during the recovery phase. Contrary to the country-level table presented in the main text, we now run these regressions with country-clustered standard errors to account for correlations among the power status of different ethnic groups in the same country. While the key result remains intact, the standard errors widen somewhat and the effect becomes insignificant earlier on.

Table E-1: Ethno-political relevance of groups in plurality-rule governments

	Dependent variable: $Pr(EGIP)_{git}$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Decline in t	0.004	0.002	0.000	-0.001	-0.002	-0.002
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)
Trough in $t-1$	-0.055**					
	(0.023)					
Trough in $t-2$ to $t-1$		-0.044*				
		(0.022)				
Trough in $t-3$ to $t-1$			-0.037*			
			(0.021)			
Trough in $t-4$ to $t-1$				-0.032		
				(0.021)		
Trough in $t-5$ to $t-1$					-0.030	
					(0.020)	
Trough in $t-6$ to $t-1$						-0.026
						(0.020)
Control sets						
Group FEs	Yes	Yes	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year Trends	Yes	Yes	Yes	Yes	Yes	Yes
Summary stats						
Adjusted-R^2	0.136	0.137	0.137	0.137	0.137	0.136
$\bar{G}$	150	150	150	150	150	150
$\bar{T}$	36.67	36.67	36.67	36.67	36.67	36.67
$\bar{G} \times \bar{T}$	5500	5500	5500	5500	5500	5500

Note(s): The table shows the results from linear group-level panel regressions of the probability of being a politically relevant group with government participation on our indicators of the timing of economic slumps. The group-level data is from the EPR-ETH 2.00 (which includes small countries missing in EPR 3.01). The dependent variable is called `status_egip` in the original data. All groups with a power status of 'junior partner' or higher are coded as unity, all others as zero. We drop countries ruled by dominant and monopoly groups as these are single-ethnicity governments which are not the object of interest in our theory. The standard errors are clustered on the country level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

F Studying the onset of slumps

This section adds a brief analysis of the timing of slump starts to the study of slump duration presented in the main text. The analysis is not designed to be exhaustive but mainly serves to highlight two points: *i)* the timing of a slump is driven by idiosyncratic factors, and *ii)* the timing is not correlated with executive constraints or the interaction proposed in this paper. The dependent variable is now a dummy indicating the first year of each slump, all preceding years are coded as zero, and all years of decline other than the first are set to missing. This approach mirrors standard practice in the literature on the onset of conflict.

Table F-1 shows the results from estimating two-way fixed effects linear probability models of slump onsets on executive constraints and a number of macroeconomic variables. Several points are worth noting. First and foremost, the level of executive constraints has no discernible effect on the probability of experiencing a slump. The coefficients are virtually zero and statistically insignificant in all columns. Second, out of the set of macroeconomic shocks, only terms of trade shocks and inflation are robust predictors of slump onsets. Changes in the commodity terms of trade, measured using data from Gruss and Kebhaj (2019), have a large effect. A 10% deterioration in the terms of trade raises the probability of a slump beginning in year t by 1.56 percentage points (t-stat $= -2.74$). The effect of inflation is more moderate. A 10% change in the acceleration of inflation leads to a 0.2 percentage point increase in the probability of falling into a slump (t-stat $= 1.87$). The level of inflation itself is not significant (not reported). Finally, the explained variation is very low, with R^2s below 5%[1], similar to the literature on growth accelerations (Hausmann et al., 2005) and episodes of sustained growth (Berg et al., 2012).

Table F-2 reports the same set of specification with an interaction of executive constraints and the level of ethnic heterogeneity—the mechanism proposed in this paper. Note that the level ethnic heterogeneity is absorbed by the country fixed effects and heterogeneity is measured according to the Ethnologue data. The estimated coefficient on the interaction term is always insignificant and, if anything, has the "wrong" sign. Models with one year lags of all right hand side variables essentially lead to the same conclusion (not reported).

Taken together, this suggests that timing of slumps is not driven by a sudden deterioration of political institutions, nor that their effect is conditional on ethnic heterogeneity. Terms of trade shocks and changes in the inflation rate continue to matter, so that external and internal macroeconomic shocks can be linked to the probability of experiencing a slump.

[1] Within R^2s are even lower but not reported here for comparison with the cross-sectional literature.

Table F-1: Onset regressions without interaction effect

	Dependent variable: $\Pr(Slump_{it} = 1 \mid Slump_{i,t-1} = 0)$					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
XCONST	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Control sets						
Terms of trade shocks	No	Yes	No	No	No	Yes
Commodity price shocks	No	No	Yes	No	No	Yes
RER undervaluation	No	No	No	Yes	No	Yes
Openness	No	No	No	Yes	No	Yes
Inflation	No	No	No	No	Yes	Yes
Summary stats						
$N \times T$	5863	4838	5287	5863	4152	3830
N	135	132	133	135	122	122
R^2	0.0397	0.0459	0.0413	0.0402	0.0553	0.0497

Note(s): The table shows the results from linear panel fixed effects regressions of the probability of a slump starting in year t conditional on it not having begun in the year before on a set of explanatory variables. All specifications include country and time fixed effects, as well as controls for population size and regional GDP. Terms of trade shocks are measured as first differences in the logarithm of the commodity terms of trade index from Gruss and Kebhaj (2019). Commodity price shocks are three aggregate commodity price indices (fuel, agriculture, and metals and minerals) from the World Bank Commodity Price Data (The Pink Sheet) and interacted with the export shares (as % of merchandise exports) in the respective sector from the World Development Indicators averaged over the whole period from 1960 to 2008. Real exchange rate undervaluation is measured using an index introduced by Rodrik (2008). Openness is measured as imports and exports over GDP using data from the Penn World Table 7.0. Inflation is measured as the first difference in the log of one plus the inflation rate from the IMF International Financial Statistics or the World Development Indicators, depending on which series is longer. The standard errors are clustered on the country level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table F-2: Onset regressions with interaction effect

Variables	Dependent variable: $\Pr(Slump_{it} = 1 \mid Slump_{i,t-1} = 0)$ (1)	(2)	(3)	(4)	(5)	(6)
$XCONST$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
$XCONST \times ELF$	0.002	0.004	0.003	0.002	0.004	0.005
	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.006)
Control sets						
Terms of trade shocks	No	Yes	No	No	No	Yes
Commodity price shocks	No	No	Yes	No	No	Yes
RER undervaluation	No	No	No	Yes	No	Yes
Openness	No	No	No	Yes	No	Yes
Inflation	No	No	No	No	Yes	Yes
Summary stats						
$N \times T$	5863	4618	5287	5863	3958	3643
N	135	132	133	135	122	121
R^2	0.0398	0.0491	0.0415	0.0403	0.0583	0.0534

Note(s): The table shows the results from linear panel fixed effects regressions of the probability of a slump starting in year t conditional on it not having begun in the year before on a set of explanatory variables. All specifications include country and time fixed effects, as well as controls for population size and regional GDP. Terms of trade shocks are measured as first differences in the logarithm of the commodity terms of trade index from Gruss and Kebhaj (2019). Commodity price shocks are three aggregate commodity price indices (fuel, agriculture, and metals and minerals) from the World Bank Commodity Price Data (The Pink Sheet) and interacted with the export shares (as % of merchandise exports) in the respective sector from the World Development Indicators averaged over the whole period from 1960 to 2008. Real exchange rate undervaluation is measured using an index introduced by Rodrik (2008). Openness is measured as imports and exports over GDP using data from the Penn World Table 7.0. Inflation is measured as the first difference in the log of one plus the inflation rate from the IMF International Financial Statistics or the World Development Indicators, depending on which series is longer. The standard errors are clustered on the country level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Additional references

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003). Fractionalization. *Journal of Economic Growth 8*(2), 155–194.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2016). Ethnic inequality. *Journal of Political Economy 124*(2), 428–488.

Alesina, A. and E. Zhuravskaya (2011). Segregation and the quality of government in a cross section of countries. *American Economic Review 101*(5), 1872–1911.

Ashraf, Q. and O. Galor (2013). The "Out of Africa" hypothesis, human genetic diversity, and comparative economic development. *American Economic Review 103*(1), 1–46.

Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory 13*, 315–352.

Berg, A., J. D. Ostry, and J. Zettelmeyer (2012). What makes growth sustained? *Journal of Development Economics 98*(2), 149–166.

Bluhm, R., D. de Crombrugghe, and A. Szirmai (2019). Do weak institutions prolong crises? On the identification, characteristics, and duration of declines during economic slumps. *World Bank Economic Review, forthcoming*.

Boockmann, B. and A. Dreher (2003). The contribution of the IMF and the World Bank to economic freedom. *European Journal of Political Economy 19*(3), 633–649.

Desmet, K., S. Weber, and I. Ortuño-Ortín (2009). Linguistic diversity and redistribution. *Journal of the European Economic Association 7*(6), 1291–1318.

Diebold, F. X. and C. Chen (1996). Testing structural stability with endogenous breakpoint a size comparison of analytic and bootstrap procedures. *Journal of Econometrics 70*(1), 221–241.

Dreher, A. (2006). IMF and economic growth: The effects of programs, loans, and compliance with conditionality. *World Development 34*(5), 769–788.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth 8*(2), 195–222.

Gruss, B. and S. Kebhaj (2019, January). Commodity terms of trade: A new database. IMF Working Papers 19/21, International Monetary Fund.

Hansen, B. E. (2000). Testing for structural change in conditional models. *Journal of Econometrics 97*(1), 93–115.

Hausmann, R., L. Pritchett, and D. Rodrik (2005). Growth accelerations. *Journal of Economic Growth 10*(4), 303–329.

Henisz, W. J. (2000). The institutional environment for economic growth. *Economics & Politics 12*(1), 1–31.

Ilzetzki, E., C. M. Reinhart, and K. S. Rogoff (2017, February). Exchange arrangements entering the 21st century: Which anchor will hold? Working Paper 23134, National

Bureau of Economic Research.

Papell, D. H. and R. Prodan (2014). Long run time series tests of constant steady-state growth. *Economic Modelling 42*, 464 – 474.

Rodrik, D. (2008). The real exchange rate and economic growth. *Brookings Papers on Economic Activity 39*(2), 365–439.