The Golden Halo and Political Transitions

Toke S. Aidt,* Facundo Albornoz† and Martin Gassebner‡

October 2009 – Preliminary version please do not quote

Keywords: political transitions; democracy; autocracy; political instability;

1 Introduction

The emergence of a new political regime is often followed by loan agreements with international organizations such us the World Bank and the IMF. This is true for both new democracies and new autocracies. For example, IMF agreements with newly established political regimes range from Bolivia in 1956 to Jordan in 1990, include emerging democracies as Spain in 1978 or Turkey in 1979, as well as emerging dictatorships like Chile in 1974 or Argentina in 1976, and involve a considerable amount of money. We refer to these capital inflows to either new democracies or new autocracies as "golden halos" and investigate their effect on political stability.

This unexplored question comes with the perk of allowing for an empirical test of Acemoglu and Robinson's theory of political transitions. In a series of papers and a subsequent book, Acemoglu and Robinson refreshed the analysis of the determinants of autocracy and democracy. This work attracted a plethora of favorable reviews and became the main reference in the field. Criticisms focus on the empirical implementation and relevance of some of the pieces of their argument. Of particular importance is that the type of political regime emerging in equilibrium depends of how costly is for the economic elite to mount a coup and for the citizens to organize a revolution, being these variables difficult to quantify.

We propose a new test of their framework. We extend their theory of political transitions by incorporating the possibility of "golden halos". More specifically, we assume that a newly established political regime after a transition to either democracy (after a process of democratization) or autocracy (after a coup) *may* receive a transfer from abroad. This feature implies the following prediction: the expectation of golden halos increases the probability of a political regime change;

*Faculty of Economics, University of Cambridge; CB3 9DD Cambridge; U.K.; toke.aidt@econ.cam.ac.uk
†Department of Economics, University of Birmingham. E-mail: f.albornoz@bham.ac.uk.
‡Corresponding author: KOF Swiss Economic Institute, ETH Zurich; Weinbergstrasse 35; CH-8092 Zurich; Switzerland; gassebner@kof.ethz.ch

that is, it causes regime instability. This is an important prediction as one the intentions behind IMF and World Bank agreements with new regimes is to achieve political stability. Moreover, golden halos may induce democratization if available exclusively to emerging democracies.

How golden halos affect the stability of newly established political regimes? Following Acemoglu and Robinson, we assume that democratic majority (the poor) imposes tougher taxation than autocracies representing the interests of the economic elite (the rich). Under autocracy, the poor can only induce redistribution through the threat of revolution. If binding enough, the elite will offer concessions but these are as temporary as the revolution threat. When temporary redistribution is not enough, democratization appears as the only credible action against the prospects of a revolution. Under democracy, the threat of a coup can temper redistribution pressure. Again, if temporary democratic concessions are not enough, a coup can restore rich citizens' hegemony through the (re) emergence of an autocratic regime. Both revolution and coup threats depend on the resources destroyed in the process of revolution or autocracy restoration. Transfers to new regimes influence the elite's incentives associated with extending voting rights (democracy is more beneficial) and mounting a coup (lower costs). If large enough, anticipations of a golden halo can trigger a regime switch that would not otherwise have taken place. If neutral with respect to whether the new regime is a democracy or an autocracy, golden halos imply regime instability. If biased in favor of a particular political regime, expected golden halos can either reduce or spread democratization around the world.

The impact of golden halos is therefore of empirical nature. Thus, we estimate the probability of a regime transition across the world between xxx and xxx and test the effect of diverse definitions of golden halos. These measures differ in two dimensions, the source of golden halos and how political agents predict their possibility. As to the source, we restrict our attention to different level of agreements with the World Bank and the IMF. This is important as they involve significant resources made available to both new democracies and new dictatorships. As to predictability, we consider specifications with either rational or adaptive expectations. We approximate rational expectations by an indication of whether a golden halo was received by the country after a transition. That is, the effect of a fully anticipated golden halo. In the adaptive expectations version, we build an indicator that weights the country's past record of golden halos with the experience of the country's neighbors. This allows for different levels of information considered by political agents in the prediction of a golden halo. In a narrow specification, only national experience counts. This is not only conceptually restrictive but also generates potential bias in our estimations as golden

halos can be correlated with unobserved factors triggering regime transitions. We therefore use broader definitions that include the experience of neighbor countries. Arguably, neighbor effects can be considered to be exogenous from the point of view of the country and therefore they reduce the risk of endogeneity in estimating the effect of a golden halo.

Our analysis is related to a series of literatures and open questions. There is a literature on the origins and dynamics of different political regimes and institutions in addition to AR's work (Boix, 2003; Gradestein, 2007; Lizzeri and Perico, 2004). This literature has mainly focused on domestic factors such as inequality, growth volatility and economic development. Far less emphasis has been placed on the role played by the international community. Exception are Acemoglu and Robinson (2005, chapter 10) and Boix (2003). While AR focus on the effects of international trade, Boix argues that capital mobility reduces redistribution under democracies which in turns facilitates democratization. In a previous paper, we analyze the effect of foreign countries in sponsoring coups, stabilizing dictatorships and facilitating constrained democratization (AA 2009). Easterly, Satyanath and Berger (2008) provide evidence of US and Soviet interventions and quantify their impact as a decline in democracy across the world of about 33%.

The literature on the empirical determinants of democracy and autocracy is large and offer a great variety of potential determinants. Gassebner, Lamla and Vreeland (2009) find that, of among 59 factors, GDP per capita and past transitions are the most robust determinants of the establishment and consolidation of democracy. To our knowledge, no previous study has investigated theoretically or empirically the effects of transfers to new political regimes. We therefore offer a novel determinant that complements previous analysis.

The effects of the IMF and to a lesser extent of the World Bank have also attracted a lot of work. Vreeland (2003) who finds that IMF agreements reduce economic growth and increase inequality. More recently, Barro and Lee (2005) also find a negative effect on economic growth, but add that IMF loans increase trade openness and reduce both the rule of law and democracy. The antidemocratic effect of the IMF goes in favor of our argument although we focus on a well defined time window of the agreement.

2 The Model

In this section, we outline and extend the theory of political transition proposed by Acemoglu and Robinson (2001). We have simplified the theory in several dimensions. These simplifications facilitate the exposition but are not critical for the point we want to make. We consider a society

with infinite time horizon, $t = 0, 1, ...\infty$. Incomes are discounted by the factor β. It is populated by two groups of individuals, the rich and poor. The total size of the population is normalized to 1 and the fraction of poor is $\lambda > \frac{1}{2}$. The political regime (S_t^{Pol}) of the society can be either democracy ($\mathcal{D}$), autocracy ($\mathcal{A}$) or socialism ($\mathcal{S}$), i.e., the *political state* is $S_t^{Pol} \in \{\mathcal{D}, \mathcal{A}, \mathcal{S}\}$. Regime transitions happen through coups, revolutions, or democratization. The opportunities for coups and revolutions depend on many different political, technological and economic factors. To capture this, we assume that the costs of coups and revolutions are stochastic and depend on the *social state* ($S_t^s \in \{G, B\}$). When the social state is G, conditions for either a coup or a revolution are favorable and the costs are relatively low (see below). When the social state is B, a coup or a revolution is prohibitively costly. The probability that the social state is G (B) is denoted ψ ($1 - \psi$).[1]

We specify the per-period incomes of the members of the two groups directly as functions of the political states and denote them by $y_i\left(S_t^{Pol}\right)$ for $i \in \{R, P\}$.[2] Utility is linear in incomes. Under autocracy, the rich controls the government and no redistribution takes place. The income of the rich is $y_R(\mathcal{A})$ while that of poor is $y_P(\mathcal{A}) < y_R(\mathcal{A})$. Under democracy the poor holds the majority and use the state to redistribute income from the rich. As a consequence, $y_R(\mathcal{A}) > y_R(\mathcal{D}) > 0$ and $y_P(\mathcal{A}) < y_P(\mathcal{D})$. Finally, under socialism wholesale expropriation of the rich takes place and we assume that $y_R(\mathcal{S}) = 0$ and $y_P(\mathcal{S}) > y_P(\mathcal{D})$.

The poor might initiate a revolution to change the political state from autocracy to socialism. We assume that socialism is an absorbing state. During a revolution, however, some income, $\mu_{S_t^s}$, is lost. How much depends on the social state. If $S_t^s = B$, then $\mu_B = \infty$ and the poor never attempt a revolution. If, on the other hand, $S_t^s = G$, then $\mu_G = \mu < \infty$ and they might be willing to pay the price of a revolution.

The rich have a strong incentive to avoid a revolution because they lose everything. The only way to avoid a revolution is to give the poor the right to vote. This leads to a transition to democracy, as we assume throughout that the poor prefer any type of democracy to socialism. A sufficient condition is that $\mu > \underline{\mu}$ where[3]

$$\underline{\mu} \equiv \frac{y_P(\mathcal{S}) - y_P(\mathcal{D})}{1 - \beta} + \frac{\beta\psi\left(y_P(\mathcal{D}) - y_P(\mathcal{A})\right)}{\left(1 - \beta(1 - 2\psi)\right)(1 - \beta)}. \tag{1}$$

[1] Acemoglu and Robinson (2001) link, for concreteness, the conditions for social unrest directly to the business cycle. In fact, they assume that coups and revolutions can only take place during recessions. We prefer to focus on (exogenous) political factors.

[2] These incomes can be derived from more fundamental assumptions about endowments, production technologies and tax instruments as in Acemoglu and Robinson (2001). Doing so complicates the analysis without affecting our main results.

[3] We derive this condition in Appendix.

Such a transition may, however, be temporary only: the rich can namely mount a coup to reinstate autocracy. A coup is costly because of the turmoil it creates. As a consequence, some of the income of the rich, $\phi_{S_t^s}$, is lost during a coup. How much again depends on the social state. If $S_t^s = B$, then $\phi_B = \infty$ and the rich never attempt a coup. If, on the other hand, $S_t^s = G$, then $\phi_G = \phi < \infty$ and the rich might be willing to pay the price of a coup.

The new feature of the model is the "golden halo". Specifically, we assume that a newly established political regime after a transition to either democracy (after a process of demomcratization) or autocracy (after a coup) *may* receive a one-off gift or transfer from abroad.[4] We assume that the transfer is distributed equally across the population and denote the per-capita transfers by $\widehat{\sigma}_j \geq 0$ with $j \in \{\mathcal{A}, \mathcal{D}\}$.[5] The size of the golden halo is unknown before the transition. We assume that it is drawn from a stationary distribution with mean σ_j and variance υ_j. The draw takes place immediately after each transition and is independent of past draws. The presence of a golden halo affects, as we shall see, regime dynamics in interesting and surprising ways and provides a prediction of the theory that we can test directly.

The timing of events within each period is as follows:

1. The social state $S_t^s \in \{G, B\}$ is revealed.

2. If a revolution has happened in the past, then the political regime is socialism and the period ends and incomes are $y_i(\mathcal{S})$ for $i \in \{R, P\}$.

3. If $S_t^{Pol} = \mathcal{A}$, the rich may democratize. If $S_t^{Pol} = \mathcal{D}$, the rich may initiate a coup that leads to autocracy. If a political transition takes place, incomes are determined by the new regime; otherwise they are determined by the old regime. Another regime transition cannot happen within that period.

4. If $S_t^{Pol} = \mathcal{A}$, the poor can initiate a revolution which leads to socialism. If no revolution takes place, incomes are realized as described by stage 2 or 3.

5. Incomes are consumed and the period ends. If a political transition to either $\mathcal{A}$ or $\mathcal{D}$ happened within the period, the size of the golden halo is realised and transfer is distributed among the population.

[4] Logically, there is a third possibility, namely that a socialistic regime (after a revolution) receives a transfer. Although this might have been important during the Cold War, we do not consider this in the present paper. We believe that the analysis of transitions to socialism is an important topic that deserves attention, but it goes beyond the scope of the present paper to provide a proper analysis.

[5] This is a simplifying assumption that can be modified. Our results hold as long as the rich benefit from the golden halo.

We treat the members of the two groups as two players of a dynamic game. We restrict attention to pure strategy Markov perfect equilibria (MPEs). A Markov perfect strategy determines for each player the appropriate action as a function of the current state of the world only, i.e., $\left(S^S, \mathcal{A}\right)$, $(S^S, \mathcal{D})$ or $\mathcal{S}$ where $S^S \in \{G, B\}$. In state $(S^S, \mathcal{A})$, the action space of the rich consists of a decision to democratize or not, while in state $(S^S, \mathcal{D})$, the action space of the elite is to mount a coup or not. Since state $\mathcal{S}$ is absorbing, we need not specify the strategy of the rich in this state. When the state is $(S^S, \mathcal{A})$, a strategy of the poor is a function of the state of the world and the rich's decision to democratize or not. When the state is $(S^S, \mathcal{D})$, poor's strategy is simply a function of the state. The strategy determines the appropriate action of the poor. In state $(S^S, \mathcal{A})$, their action space is a decision to mount a revolution or not, while in state $(S^S, \mathcal{D})$, they are not required to take any actions. A pure strategy Markov perfect equilibrium is then defined as a set of strategies for rich and the poor that are best responses to each other for all possible states.

3 Analysis and Results

We assume that the initial political state is autocracy. The effect of a golden halo on regime dynamics and stability depends critically on whether the poor can credibly threaten to organize a revolution to overthrow the autocracy or not. The decision to organize a revolution is made at stage 4 of the game. It is based on the following considerations. If a revolution is organized, the outcome is socialism for ever and the poor get $\frac{y_P(\mathcal{S})}{1-\beta} - \mu_{S_t^s}$. It is clear that they have no incentive to organize a revolution in social state B (as $\mu_{S_t^s} = \infty$). In social state G, on the other hand, they might organize a revolution, but it depends on how badly the poor fare under autocracy. Under (perpetual) autocracy, the poor get $\frac{y_P(\mathcal{A})}{1-\beta}$. Therefore, the poor never organize a revolution in state $(G, \mathcal{A})$ when

$$\mu \geq \mu^* \equiv \frac{y_P(\mathcal{S}) - y_P(\mathcal{A})}{1-\beta}. \tag{2}$$

When this so-called revolution constraint in binding, i.e., $\mu < \mu^*$, the rich must democratize to avoid socialism.[6] This leads to democracy. Importantly, however, the golden halo opens another path to democracy that applies even when the cost of revolution in state G is so large that the the poor never attempt a revolution ($\mu > \mu^*$). It is possible that the rich might hand over power to the poor just to trigger the golden halo!

[6] Note that $\mu^* > \underline{\mu}$.

The revolution constraint is not binding We begin the equilibrium characterization by considering the case in which the revolution constraint never binds ($\mu > \mu^*$). The poor finds it too expensive to organize a revolution whatever the social conditions are: they prefer perpetual autocracy to a revolutionary transition to socialism. In this case, any transition to democracy is voluntary but the transition is not inevitable and may not last. When the transition to democracy is for good, we say that the economy transits to *perpetual democracy.* On the other hand, when the transition to democracy is only temporary, we say that the economy transits to *unstable democracy.* In the latter case, the rich grant voting rights to the poor in the very first period, but mount a coup against the democracy at the next opportunity, for again to grant voting rights after just one period of autocracy. Finally, if no political transitions ever take place, we say that the economy is a *perpetual autocracy.*

Since by assumption $\mu > \mu^*$, the poor never attempt a revolution at stage 4. Anticipating that at stage 3, the rich effectively face the choice between three strategies:

1. Perpetual autocracy: Irrespective of the social state, the rich never democratize. The economy continues to be autocratic and the rich get $\frac{y_R(\mathcal{A})}{1-\beta}$.

2. Perpetual democracy: Irrespective of the social state, the rich democratize in the first period and never attempt a coup in subsequent periods. The economy is a democracy for ever and the rich expect to get $\frac{y_R(\mathcal{D})}{1-\beta} + \sigma_{\mathcal{D}}$ where $\sigma_{\mathcal{D}}$ is the expected value of the golden halo after a democratization.

3. Unstable democracy: Irrespective of the social state, the rich democratize each time the political state is $\mathcal{A}$ and initiate a coup each time the state is $(G, \mathcal{D})$.[7] The rich expect to get[8]

$$\frac{y_R(\mathcal{D}) + \psi\beta y_R(\mathcal{A}) + (1 - (1-\psi)\beta)\sigma_{\mathcal{D}} + \psi\beta(\sigma_{\mathcal{A}} - \phi)}{(1-\beta)(1+\psi\beta)} \tag{3}$$

   where $\sigma_{\mathcal{D}}$ and $\sigma_{\mathcal{A}}$ are the expected values of the golden halo after a transition to democracy and autocracy, respectively.

The equilibrium strategy of the rich depends on the value of $\sigma_{\mathcal{D}}$, $\sigma_{\mathcal{A}}$ and ϕ. We can defined the following three threshold values. Firstly, a direct comparison between strategy 1 and 2 shows

[7] If democratization should be followed by a coup, it is never optimal for the elite to democratize and then not to initiate a coup the first time after that $S_t^S = G$. Thus, we can focus on the comparison of strategy 2 and strategy 3.

[8] See Appendix for details.

that the rich prefer perpetual democracy to perpetual autocracy if and only if $\sigma_{\mathcal{D}} > \sigma^1_{\mathcal{D}}$ where

$$\sigma^1_{\mathcal{D}} = \frac{y_R(\mathcal{A}) - y_R(\mathcal{D})}{1-\beta}. \tag{4}$$

Secondly, a comparison between strategy 1 and 3 shows that the rich prefer unstable democracy to perpetual autocracy if and only if $\sigma_{\mathcal{D}} > \sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})$ where

$$\sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}}) = \frac{y_R(\mathcal{A}) - y_R(\mathcal{D})}{1-(1-\psi)\beta} + \frac{\psi\beta(\phi - \sigma_A)}{1-(1-\psi)\beta}. \tag{5}$$

Thirdly, comparing strategies 2 and 3, we see that the rich prefer unstable democracy to perpetual democracy if and only if $\sigma_{\mathcal{D}} > \sigma^3_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})$ where

$$\sigma^3_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}}) = -\frac{y_R(\mathcal{A}) - y_R(\mathcal{D})}{\beta} + \frac{(\phi - \sigma_A)}{\beta}. \tag{6}$$

Given these thresholds, we can state the following result.

Proposition 1 *Suppose the initial political state is autocracy and that $\mu > \mu^*$. Then for all $\sigma_{\mathcal{D}} \neq \left\{\sigma^1_{\mathcal{D}}, \sigma^2_{\mathcal{D}}, \sigma^3_{\mathcal{D}}\right\}$ there exists a unique pure strategy MPE such that*

1. *If $\sigma_{\mathcal{D}} > \max\left\{\sigma^2_{\mathcal{D}}, \sigma^3_{\mathcal{D}}\right\}$ then the economy becomes an unstable democracy. The rich democratize each time the political state is $\mathcal{A}$ and mount a coup each time the state is $(G, \mathcal{D})$.*

2. If $\sigma_{\mathcal{D}} > \sigma^1_{\mathcal{D}}$ and $\sigma_{\mathcal{D}} < \sigma^3_{\mathcal{D}}$, then the economy *becomes a perpetual democracy. The rich democratize in the first period and never attempt a coups after that.*

3. Otherwise, the economy is a perpetual autocracy.

Proof. Begin by noting the following facts about the three thresholds defined in the text above. There exists a value of the cost of a coup, $\widetilde{\phi}$, such that i) $\sigma^2_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}}) = \sigma^3_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}}) = \sigma^1_{\mathcal{D}}$, ii) $\sigma^1_{\mathcal{D}} \geq \sigma^2_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}}) \geq \sigma^3_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}})$ for $\phi \leq \widetilde{\phi}$ and iii) $\sigma^3_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}}) > \sigma^2_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}}) > \sigma^1_{\mathcal{D}}$ for $\phi > \widetilde{\phi}$. The optimal strategy of the poor is not to initiate a revolution ever. Given that, the decision of the rich to democratize or not is independent of the social state and the rich democratize only when it is in their interest to do so. The rich prefer unstable democracy to perpetual autocracy or democracy if and only if $\sigma_{\mathcal{D}} > \sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})$ and $\sigma_{\mathcal{D}} > \sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})$. The rich prefer perpetual democracy to perpetual autocracy or unstable democracy if and only if $\sigma_{\mathcal{D}} > \sigma^1_{\mathcal{D}}$ and $\sigma_{\mathcal{D}} < \sigma^3_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})$. The rich prefer perpetual autocracy to the other alternatives if and only if $\sigma_{\mathcal{D}} < \sigma^1_{\mathcal{D}}$ and $\sigma_{\mathcal{D}} < \sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})$.

The equilibrium strategy of the rich then is i) if $\sigma_{\mathcal{D}} > \max\left\{\sigma_{\mathcal{D}}^2, \sigma_{\mathcal{D}}^3\right\}$, democratize when the state is $(S^S, \mathcal{A})$ for $S^S \in \{B, G\}$, mount a coup when the state is $(G, \mathcal{D})$, and do nothing when the state is $(B, \mathcal{D})$; ii) if $\sigma_{\mathcal{D}} > \sigma_{\mathcal{D}}^1$ and $\sigma_{\mathcal{D}} < \sigma_{\mathcal{D}}^3(\phi, \sigma_{\mathcal{A}})$, democratize in period 1 irrespective of the social state and never attempt a coup; iii) If $\sigma_{\mathcal{D}} < \min\left\{\sigma_{\mathcal{D}}^1, \sigma_{\mathcal{D}}^2(\phi, \sigma_{\mathcal{A}})\right\}$, never democratize and never attempt a coup ■

In the absence of a credible threat of revolution ($\mu > \mu^*$) and with the average golden halo being zero ($\sigma_{\mathcal{D}} = \sigma_{\mathcal{A}} = 0$), the only possible equilibrium outcome is, as in Acemoglu and Robinson (2001), perpetual autocracy. So, expectations of a golden halo may induce democratization in situations where autocracy would otherwise have been perpetual. This is illustrated in Figure 1. What is required is that the expected value of the golden halo to a newly established democracy ($\sigma_{\mathcal{D}}$) is sufficiently large. The stability of the emerging democracy depends on the cost of a coup relative to the size of the expected value of the golden halo. For low values of ϕ (in area UD), the economy experiences repeated regime switches. For sufficiently high values of ϕ and moderately high values of σ_D (in area PD) perpetual democracy emerges. Interestingly, even if $\sigma_{\mathcal{D}} = 0$ and a newly established democracy is cannot expect to be rewarded with a golden halo, it is still possible that the rich democratize voluntarily. This happens if unstable democracy yields higher payoff than perpetual autocracy (which for $\sigma_{\mathcal{D}} = 0$ is preferred by the rich to perpetual democracy). A simple calculation shows that this requires that $\sigma_{\mathcal{A}} \geq \frac{y_R(\mathcal{A}) - y_R(\mathcal{D})}{\psi\beta} + \phi$. Thus, if the expected golden halo to a newly established autocracy is sufficiently larger, it is optimal for the rich to democratize, not because this is desirable in itself, but because of the expectation of the golden halo triggered when the rich take power back in a future coup. The area labeled PA corresponds to the equilibrium with perpetual autocracy.

The revolution constraint is binding Next, suppose that the revolution constraint binds ($\mu < \mu^*$), that is, the poor would organize a revolution whenever social conditions are favorable. In this case, the transition to democracy is inevitable: the rich will grant voting rights to avoid the transition to socialism and this is independent of the presence of the golden halo. Clearly, if the rich were willing to grant voting rights in the absence of a credible threat of revolution, they continue to show this willingness when the threat is credible. In other words, for $\sigma_{\mathcal{D}} > \max\left\{\sigma_{\mathcal{D}}^2, \sigma_{\mathcal{D}}^3\right\}$ or $\sigma_{\mathcal{D}} > \sigma_{\mathcal{D}}^1$ and $\sigma_{\mathcal{D}} < \sigma_{\mathcal{D}}^3$, the equilibrium strategies are as described in proposition 1 and the economy either transits to perpetual or unstable democracy as appropriate. So, the revolution constraint only makes a difference when the rich in the absence of a credible threat of revolution

prefer perpetual autocracy to the alternatives (i.e., when $\sigma_{\mathcal{D}} < \sigma_{\mathcal{D}}^1$ and $\sigma_{\mathcal{D}} < \sigma_{\mathcal{D}}^2(\phi, \sigma_{\mathcal{A}})$). For the rest of the section we assume that is the case.

We make a distinction between two types of democracy that might emerge: Consolidated democracy emerges when the transition is permanent. In contrast, unconsolidated democracy emerges when the transition is only temporary. In this case, the rich mount a coup at the next opportunity (in state $(G, \mathcal{D})$), for again to issue voting rights when the situation requires it (in state $(G, \mathcal{A})$).[9] Whether the democracy consolidates or not depends on the incentives of the rich to mount coups. This incentive is controlled by the so-called coup constraint. To derive this constraint suppose that the political state is democracy and let $W_i(S_t^{Pol})$ be the continuation value for group i when the political state is S_t^{Pol}. Clearly, in social state B, the rich will not mount a coup because $\phi_B = \infty$. In social state B, the situation is different. If they don't mount a coup they get $y_R(\mathcal{D}) + \beta W_R(\mathcal{D})$ and if they do, the coup triggers a golden halo to the new autocracy and they expect to get $y_R(\mathcal{A}) - \phi + \sigma_D + \beta W_R(\mathcal{A})$. The rich will never mount a coup if

$$\phi > y_R(\mathcal{A}) - y_R(\mathcal{D}) + \beta\left(W_R(\mathcal{A}) - W_R(\mathcal{D})\right) + \sigma_{\mathcal{A}}. \tag{7}$$

Since, by assumption, the current political state is democracy, it must be true that the rich were forced to democratize the last time the state was $(G, \mathcal{A})$ and that they will have to do so again next time the state is $(G, \mathcal{A})$. This is implies that the value of autocracy is

$$W_R(\mathcal{A}) = \psi\left(y_R(\mathcal{D}) + \sigma_{\mathcal{D}} + \beta W_R(\mathcal{D})\right) + (1-\psi)\left(y_R(\mathcal{A}) + \beta W_R(\mathcal{A})\right), \tag{8}$$

where we notice that the possible transition back to democracy if the social state is G in the next period triggers another golden halo with expected value $\sigma_{\mathcal{D}}$, this time to the new democracy. Combining this with the observation that $W_R(\mathcal{D}) = y_R(\mathcal{D}) + \beta W_R(\mathcal{D})$ under condition (7), we can write the coup constraint as

$$\sigma_{\mathcal{D}} < \frac{(\phi - \sigma_{\mathcal{A}})\left(1 - (1-\psi)\beta\right)}{\psi\beta} - \frac{y_R(\mathcal{A}) - y_R(\mathcal{D})}{\beta\psi} \equiv \sigma_{\mathcal{D}}^4(\phi, \sigma_{\mathcal{A}}). \tag{9}$$

The cut-off $\sigma_{\mathcal{D}}^4$ has a natural interpretation. The rich are only willing to mount a coup if it pays off. This is less likely to be case if the net expected cost of a coup $(\phi - \sigma_{\mathcal{A}})$ is high or when the

[9] Notice that consolidated democracy differs from perpetual democracy because autocracy may persist for some periods (until the first time the social state is G). Unconsolidated democracy differs from unstable democracy because a coup is followed by a period of autocracy (until the next time the social state is G) rather than by an immediate transition back to democracy.

payoff differential between democracy and autocracy, $y_R(\mathcal{A}) - y_R(\mathcal{D})$, is small.

The next proposition characterizes equilibrium outcomes for the case with a binding revolution constraint and in which the rich prefer perpetual autocracy to any form of voluntary democracy.

Proposition 2 *Suppose the initial political state is autocracy. Furthermore, assume that $\mu < \mu^*$ and $\sigma_{\mathcal{D}} < \min\left\{\sigma^1_{\mathcal{D}}, \sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})\right\}$. Then for all $\sigma_{\mathcal{D}} \neq \sigma^4_{\mathcal{D}}$ there exists a unique pure strategy MPE such that*

1. *If $\sigma_{\mathcal{D}} < \sigma^4_{\mathcal{D}}$, then the economy becomes a consolidated democracy. The rich democratize the first time the social state is G and never attempts a coups after that.*

2. *If $\sigma_{\mathcal{D}} > \sigma^4_{\mathcal{D}}$, then the economy becomes an unconsolidated democracy. The rich democratize each time the state is $(G, \mathcal{A})$ and mount a coup each time the state is $(G, \mathcal{D})$.*

Proof. The initial political state is $\mathcal{A}$. In autocracy, the poor moves after the rich. In state $(B, \mathcal{A})$, the best response of the poor no matter what the rich do is not to oreganize a revolution. Anticipating that, the elite does not democratize (as $\sigma_{\mathcal{D}} < \min\left\{\sigma^1_{\mathcal{D}}, \sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})\right\}$). In state $(G, \mathcal{A})$, the poor will organize a revolution if the rich do not democratize. Anticipating this, the best response of the rich is to democratize. In state $(B, \mathcal{D})$, the poor do to not make any choice. The rich will not mount a coup because the cost of doing so is infinite. In state $(G, \mathcal{D})$, the poor do not make any choice. The rich will mount a coup if $\sigma_{\mathcal{D}} > \sigma^4_{\mathcal{D}}$ and not mount one otherwise. To complete the proof, we need to show that both cases are consistent with $\sigma_{\mathcal{D}} < \min\left\{\sigma^1_{\mathcal{D}}, \sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})\right\}$. We notice that there exist a unique $\widetilde{\phi} = \frac{y_R(\mathcal{A}) - y_R(\mathcal{D})}{1-\beta} + \sigma_A$ such that $\sigma^2_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}}) = \sigma^4_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}}) = \sigma^1_{\mathcal{D}}$. Moreover, at $\phi = 0$, $\sigma^4_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}}) < \sigma^2_{\mathcal{D}}(\widetilde{\phi}, \sigma_{\mathcal{A}})$. This implies that for $\phi \in [0, \widetilde{\phi})$ there exist values of $\sigma_{\mathcal{D}}$ such that $\sigma_{\mathcal{D}} < \sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})$ and $\sigma_{\mathcal{D}} > \sigma^4_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})$ and that there exist values $\sigma_{\mathcal{D}}$ such that $\sigma_{\mathcal{D}} < \sigma^4_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}}) < \sigma^2_{\mathcal{D}}(\phi, \sigma_{\mathcal{A}})$ ■

Intuitively, the proposition shows that consolidated democracy emerges when the cost of a coup is high, while unconsolidated democracy with frequent regime changes arises when the cost is sufficiently low. This is illustrated in Figure 2. Area CD corresponds to consolidated democracy and area UCD corresponds to unconsolidated democracy. Importantly, regime dynamics is affected directly be the presence of the golden halo. In particular, we have:

Proposition 3 *(Golden Hallo) An increase in the expected value of the golden halo increases regime instability by making a transition to unconsolidated democracy more likely.*

Proof. The proposition follows from the fact that unconsolidated democracy is more likely when $\sigma_{\mathcal{D}}$ is large and that $\frac{\partial \sigma_{\mathcal{D}}^4}{\partial \sigma_{\mathcal{A}}} = \frac{1-\beta(1-\psi)}{\beta\psi} > 0$ ■

The proposition shows that the expectation of a golden halo may cause an otherwise consolidated democracy to become unconsolidated and induce regime volatility. The intuition for this result is straight forward: regime volatility triggers frequent golden halos.

4 An Empirical Test

The testable implication of the extended version of Acemoglu and Robinson's (2001) theory of political transition is clear-cut: the expectation of a golden halo increases the probability of a regime change. It causes regime instability. Exogenous variations in the expected value of the golden halo causes exogenous shifts in the coup constraint that can help us identify the impact on the regime transition probability and/or on regime duration.

4.1 The Testing Strategy

More specifically, according to the theory the decision to democratize or to overthrow an existing democracy is affected by *expectations* about a golden halo. In other words,

$$\Pr(PT_{it} = 1) = F(GH^e_{it-1}; X_{it})$$

where PT is equal to one if a regime transition takes place at time t in country i, and X_{it} is a vector of observable factors that affect the probability of a regime transition. The key variable is GH^e_{it-1}. It represents the expectation formed at time $t-1$ that the country will receive a golden halo at time t if a political transition takes place. The theoretical prediction is that $\frac{\partial F}{\partial GH^e_{it-1}} > 0$.

We measure GH^e_{it-1} in two alternative ways. Firstly, suppose that the political actors form rational expectations such that they given all available information, on average, get it right. Empirically, we approximate the rational expectation by the lead of an indicator variable that takes the value 1 if a golden halo was received after a transition. In other words, we ask if a country that perfectly anticipates getting a golden halo is more likely to undergo a transition than a country that (correctly) anticipates not getting one.

Secondly, we suppose that the political actors use an adoptive strategy to estimate the likelihood that a golden halo will be forthcoming. In this case, they will be looking at the country's own past experience and/or at the experience of their neighbors to estimate the likelihood that a

regime transition will trigger a golden halo. To implement this, we define a neighborhood N and calculate GH^e_{it-1} for each year as the weighted sum of all past golden halos in that neighborhood as follows:

$$GH^e_{it-1} = \sum_{i=1}^{N} \sum_{\tau=1}^{t-1} \left(x_{-i,t-1} + \lambda^{-\tau} x_{i,t-1-\tau} \right) \tag{10}$$

where x_{it} is 1 if country $i \in N$ in year $\tau \leq t-1$ get a golden halo and zero otherwise. λ is a weighting parameter. The idea is that golden halos in the more distant past carry less weight. Depending on the definition of N, equation (10) encompasses three special cases that we shall make use of in the estimations. One specification is to restrict the neighborhood to the country itself. This is very restrictive in terms of the information that political agents use to predict if a golden halo is likely to be forthcoming. It is also possible that golden halos in the past are correlated with unobserved factors that also affect political transitions in that country. If so, this will bias the inference. This concern motivates our two other specifications. In our second specification, we defined the neighborhood as the rest of the world, excluding the country itself. This is a very large neighborhood and our third specification reduces the size of the neighborhood to the region in which the country is located (WHICH ONES).

To implement this test, we need to define what we mean by a golden halo and what we mean by a regime transition. We use two definitions of the golden halo:

1. **Narrow definition**: a new growth-enhancing aid agreement/programme from either the IMF or the World Bank within a two years window after a regime transition.

2. **Broader definition**: any new IMF or World Bank agreement/programme (including loans) within a two years window after a regime transition.

The vector X of control variables is chosen carefully. Gassebner, Lamla and Vreeland (2009) have undertaken an extreme bounds analysis to establish which of the many potential determinants of regime transition are robust. We use the variables that they find to be robustly related to regime transitions as out baseline and add our golden halo variables to this specification.

- Construct halo variables:

  - For a given neighborhood N which can be the world or a region, we calculate for each year the weighted sum of all past golden halo in that neighborhood as follows:

$$G_{it} = \sum_{i=1}^{N} \sum_{\tau=1}^{t} \left(x_{-it} + \lambda^{-\tau} x_{it-\tau} \right)$$

where x_{it} is 1 if a golden halo happened in country i in year t and zero otherwise. λ is a weighting parameter. The idea is that golden halos in the more distant past carry less weight. Exclude own current GH.

–

- Specifications to estimate.
  - Perfectly anticipated GHs: include all programs/agreements with leads of one or two periods in a given class in the estimation.
  - Adaptively expected GH: include the various definitions of the backward looking GH variables.
  - Transitions from democracy to autocracy for subsample with countries with a previous GH.

5 Appendix

Deriving condition (3) The poor benefit from a transition to democracy for two reasons. Firstly, their income is higher than under autocracy (but lower than under socialism). Secondly, they share in the golden halo or halos if multiple transitions take place. We are seeking a condition that ensures that the poor will "cancel" the revolution if the rich grant them voting rights. In the absence of the golden halo, unconsolidated democracy, understood as a situation in which the rich grant voting rights when the state is $(G, \mathcal{A})$ and mount a coup when the state is $(G, \mathcal{D})$, defines a lower bound on the welfare of the poor under democracy. Thus, if this can prevent a revolution by dominating a transition to socialism for $\sigma_{\mathcal{D}} = \sigma_{\mathcal{A}} = 0$, so can any other type of democracy with or without a golden halo. Formally, we seek a condition that ensures

$$\frac{y_P(\mathcal{S})}{1-\beta} - \mu \leq y_P(\mathcal{D}) + \sigma_{\mathcal{D}} + \beta W_P(\mathcal{D}) \tag{11}$$

where

$$W_P(\mathcal{D}) = \psi \left(y_P(\mathcal{A}) + \sigma_{\mathcal{A}} + \beta W_P(\mathcal{A}) \right) + (1-\psi) \left(y_P(\mathcal{D}) + \beta W_P(\mathcal{D}) \right)) \tag{12}$$

and

$$W_P(\mathcal{A}) = \psi \left(y_P(\mathcal{D}) + \sigma_{\mathcal{D}} + \beta W_P(\mathcal{D}) \right) + (1-\psi) \left(y_P(\mathcal{A}) + \beta W_P(\mathcal{A}) \right). \tag{13}$$

This yields two equations in two unknown, which we can solve to get

$$W_P(\mathcal{D}) = \frac{\psi y_P(\mathcal{A}) + (1-\beta(1-2\psi)-\psi) y_P(\mathcal{D}) + \beta\psi^2\sigma_{\mathcal{D}} + (1-\beta(1-\psi))\psi\sigma_{\mathcal{A}}}{(1-\beta(1-2\psi))(1-\beta)} \quad (14)$$

$$W_P(\mathcal{A}) = \frac{\psi y_P(\mathcal{D}) + (1-\beta(1-2\psi)-\psi) y_P(\mathcal{A}) + \beta\psi^2\sigma_{\mathcal{A}} + (1-\beta(1-\psi))\psi\sigma_{\mathcal{D}}}{(1-\beta(1-2\psi))(1-\beta)}. \quad (15)$$

For $\sigma_{\mathcal{D}} = \sigma_{\mathcal{A}} = 0$, substitution of this into equation (11) and rearrange gives

$$\mu \geq \frac{y_P(\mathcal{S}) - y_P(\mathcal{D})}{1-\beta} - \frac{\beta\psi\left(y_P(\mathcal{A}) - y_P(\mathcal{D})\right)}{(1-\beta(1-2\psi))(1-\beta)} \equiv \underline{\mu}. \quad (16)$$

This is a condition that only depends on the parameters of the model, not on the strategies of the elite and workers and it is sufficient, not necessary, to prevent a revolution.

Deriving condition (3) We want to calculate the value of following strategy 3 starting from $S_t^{Pol} = \mathcal{A}$. Since the rich democratize no matter what the social state is, the value is

$$W_R(\mathcal{A}) = y_R(\mathcal{D}) + \sigma_{\mathcal{D}} + \beta W_R(\mathcal{D}). \quad (17)$$

To evaluate this, we need to calculate the continuation value starting from $S_t^{Pol} = \mathcal{D}$, i.e., $W_R(\mathcal{D})$. If the social state is G the rich mount a coup and there is a transition to autocracy and if the social state is B, the rich does nothing and the democracy persist for another period. We can therefore write

$$W_R(\mathcal{D}) = \psi\left(y_R(\mathcal{A}) - \phi + \sigma_{\mathcal{A}} + \beta W_R(\mathcal{A})\right) + (1-\psi)\left(y_R(\mathcal{D}) + \beta W_R(\mathcal{D})\right). \quad (18)$$

Solving this equation for $W_R(\mathcal{D})$ gives

$$W_R(\mathcal{D}) = \frac{\psi\left(y_R(\mathcal{A}) - \phi + \sigma_{\mathcal{A}} + \beta W_R(\mathcal{A})\right) + (1-\psi) y_R(\mathcal{D})}{1-\beta(1-\psi)} \quad (19)$$

Substituting this back into equation (17) and rearrange gives equation (3)