On the Joint Evolution of Culture and Political Institutions: Elites and Civil Society

Alberto Bisin

New York University, National Bureau of Economic Research, and Center for Economic Policy Research

Thierry Verdier

Paris School of Economics, École des Ponts ParisTech, Pontifical Catholic University of Rio de Janeiro, and Center for Economic Policy Research

We provide an abstract model of the interaction between culture and political institutions. The model is designed to study the political economy of elites and civil society on the determination of long-run socioeconomic activity. We characterize conditions such that the cultural traits of elites and civil society and the institutions determining their relative political power complement (substitute) each other, giving rise to a multiplier effect that amplifies (dampens) their combined ability to spur socioeconomic activity. We show how the joint dynamics may display hysteresis and oscillations, depending on the form of the interaction between elites and civil society.

I. Introduction

> *Era questo un ordine buono, quando i cittadini erano buoni . . . ma diventati i cittadini cattivi, divento' tale ordine pessimo.* [This was a

Thanks to Alberto Alesina, Gary Becker, Jess Benhabib, Emeline Bezin, Georgy Egorov, Martin Fiszbein, Luigi Guiso, David Levine, Simone Meraglia, Dilip Mookherjee, Massimo Morelli, Nicola Persico, Louis Putterman, Kenneth Shepsle, Shlomo Weber, and many

Electronically published April 23, 2024

Journal of Political Economy, volume 132, number 5, May 2024.

https://doi.org/10.1086/727711

> good institutional order when citizens were good . . . but when citizen became bad, it turned into an horrible order.] (Niccoló Machiavelli, *Discorsi,* I. 16, 1531 [our translation])

> Among a people generally corrupt, liberty cannot long exist. (Edmund Burke, Letter to the Sheriffs of Bristol, April 3, 1777)

> Is there no virtue among us? If there be not, we are in a wretched situation. No theoretical checks—no form of government can render us secure. To suppose that any form of government will secure liberty or happiness without any virtue in the people, is a chimerical idea. (James Madison, June 20, 1788, *Papers* 11:163)

A thriving literature in economics and political science studies which factors may account for long-run income inequality across countries in the world. In this context, institutions and culture are often run against each other as possible explanatory factors. In fact, causal effects are arguably multifaceted, depending, for example, on the time frame of reference. For instance, on the one hand, a fundamental dimension of culture, social capital, is "making democracy work" in Putnam's Italy (Putnam 1993). But on the other hand, social capital formation in Western Europe is partly determined by the historical pattern of political institutions from 1600 to 1850 (Tabellini 2008a, 2010).[1]

More generally, important historical narratives suggest that the interactions of culture and institutions have an important role in providing an understanding of the determinants of long-run economic activity and, more generally, socioeconomic prosperity. Indeed, several important historical processes are interpreted as the outcome of the interactions between the cultural traits of elites and civil society and the institutions determining their relative political power. For instance, the formation of inclusive institutions protecting property rights in England has arguably spurred economic activity after/concurrently with the spread across the elites of an appropriate system of ideas and beliefs, the "bourgeois ideology" in McCloskey (2006, 2010, 2017) and the "industrial enlightment" in

others for comments. Thanks also to the editor, Emir Kamenica, and the referees for exceptionally insightful reports. We have presented the paper widely before and after writing it, and we thank all the audiences for help, suggestions, and support. Earlier drafts of this paper circulated under the title "On the Joint Evolution of Culture and Institutions" and "On the Joint Evolution of Culture and Institutions: Elites and Civil Society." This research has been undertaken with financial support from the European Research Council advanced grant 324004 for the project TECTACOM (The Economics of Cultural Transmission and Applications to Communities, Organizations, and Markets).

[1] Furthermore, the effects of social capital may be context specific; in some cases, specific forms of social capital may in fact impede the working of democracy, such as, e.g., norms of reciprocity favoring vote buying in Paraguay (Finan and Schechter 2012).

Mokyr (2016). This system of ideas fundamentally contributed to laying the grounds for the advent of the Industrial Revolution (see also Doepke and Zilibotti 2008). Relatedly, the institutional independence from the Holy Roman Empire obtained by several Italian cities (communes) in the Middle Ages has had significant and very persistent effects on their economic prosperity, arguably also through the development of the stock of civic capital of their citizens, transmitted across generations and acquired by those who moved into these cities over the centuries (Guiso, Sapienza, and Zingales 2008, 2016).[2]

Abstracting from specific contextual instances, in this paper we provide a modeling framework to analyze the role of the interaction of culture and institutions as determinants of the political economy of elites and civil society. The stylized dynamics we obtain as a result map into novel tools for the empirical analysis of the effects of exogenous variations in cultural and/or institutional phenomena. Furthermore, our analysis illustrates various regularities occurring in different socioeconomic contexts through economic history, such as, for example, the formation and circulation of the elites, the transition away from extractive institutions, the accumulation of civic capital, and their effects on economic activity. More generally, our analysis can be interpreted to shed some light on the relationship between democracy and economic activity. Recent political events—from post-Soviet Russia to Iraq and Afghanistan—have shown how difficult it may be to initiate sustainable democratization processes leading to economic growth. In this respect, we highlight the role of culture in mediating democratization processes, identifying various specific mechanisms.[3] We show how a push toward democratization might or might not be sustainable and society might either converge to more inclusive (democratic) institutions and socioeconomic prosperity or else—depending on the historical initial distribution of cultural traits and political power of the elites—revert to extractive autocratic institutions and limited economic prosperity (sec. IV.A). We also show how, in a different context, a loss of political power of the elites (democratization) might induce

[2] Along similar lines, the Roman Empire's reliance on slavery institutions—which developed jointly and alongside its aristocracy's ideological stigma against manual occupations—crucially contributed to the fall of the empire in the fourth century CE (Schiavone [1996] 2020). Economic activity after the end of colonial institutions in Africa also appears to have been modulated by the development of cultural traits and norms of behavior originated during or before colonial times (see Lowes and Monteiro [2017] on the legacy of the rubber concessions in the Congo Free State and Lowes et al. [2017] on the Kuba Kingdom). Related references include Gorodnichenko and Roland (2011, 2017) on market institutions and individualism, Greif and Tabellini (2010, 2017) on norms of kinship and moral systems, and Boranbay and Guerriero (2019) on inclusive institutions and culture of cooperation.

[3] A controversial position for the relevance of culture in the determination of long-run economic activity is in the work of Samuel Huntington (see, e.g., Harrison and Huntington 2001). See also Landes (1985) and Diamond (2005) for broad historical perspectives on the role of culture on development.

endogenously a decline of civic capital in society, undermining its impact on economic activity (see sec. IV.B).[4]

More in detail, the modeling of the joint evolution of culture and institutions we provide is postulated on a society populated by two distinct political groups of agents—say, civil society and elites—characterized by distinct economic resources (e.g., elites have more resources or a different technology to obtain them), political power, and cultural traits. Each time period, a policy game is played between individual agents and a socioeconomic policy maker (the government). Institutions represent the relative political power of these groups in civil society to affect policy decisions. Culture represents the distribution of values and preferences within (culturally heterogeneous) groups over policy decisions in society. The government's choice maximizes a social welfare function that encodes the distribution of political power between the groups (institutions), given their preferences and values (culture). A set of government policies and agents' actions arise as societal equilibrium outcomes.

Institutions evolve as the result of a process of optimal political delegation, changing the distribution of political power to internalize externalities, lack of commitment, and other distortions leading to an inefficient societal equilibrium outcomes. As a consequence, residual decision rights over public policy tend to be delegated to those political groups that are better able (or have the highest incentives) to internalize the externalities affecting the policy game. Culture evolves over time, following socialization and cultural transmission processes whose incentives are in turn affected by equilibrium outcomes of the policy game played in society. The interdependence between institutions and culture is the fundamental factor determining their joint dynamics and their effects on economic activity, for example, on long-run economic growth.

In such a setting, we characterize the cultural and institutional dynamics of the socioeconomic system. From a normative perspective, we show that even though institutional change is designed to respond to the inefficiencies of equilibrium outcomes, the societal equilibrium at the stationary state of the dynamics is not necessarily efficient, and we characterize the determinants of the welfare properties of the dynamics. From a positive perspective, we characterize conditions under which cultural and institutional dynamics are complements—so that, for example, any institutional change that spurs economic activity is reinforced by the dynamics of culture and institutions—and conditions under which, on the contrary,

[4] The interaction of cultural and institutional factors could also shed some light on the other possible direction of causality between income and institutions: Acemoglu et al. (2008) have documented in fact how income fails to effectively explain the relative incidence of democracy in the world, as would be predicted by Lipset modernization theory (Lipset 1959).

culture and institutions are substitutes, hence weakening the effects of institutional change. We show how examples of complex dynamics can emerge from the interaction between cultural and institutional change and how these qualitative dynamics depend on whether culture and institutions are complements or substitutes. These dynamical systems will generally tend to display forms of hysteresis on the equilibrium path as well as multiple stationary states and dependence on initial conditions (lack of ergodicity). We characterize a sufficient condition to rule out limit cycles, and we show that (i) local stability requires a bound on the strength of complementarity between culture and institutions but (ii) oscillatory convergent paths may occur only when culture and institutions interact as substitutes.

We define the *cultural multiplier* as the ratio of the total effect of institutional change to its direct effect, that is, the counterfactual effect that would have occurred had the distribution of cultural traits in the population remained constant after the institutional change. Similarly, the *institutional multiplier* is defined as the ratio of the total effect of cultural change to its direct effect. We show that these multipliers have the same sign, which depends on whether culture and institutions act as complements or substitutes, positive in the first case and negative in the second. The cultural and institutional multipliers are conceptual constructs defined to distinguish (and potentially measure) the relative contribution of culture and institutions to an economic outcome of interest—for example, long-term economic growth—independent of the initial causal forcing variable. This analysis can complement, in ways we identify and illustrate, the recent wave of causal analyses of either institutions or culture on future socioeconomic prosperity in historical economics and persistence studies.[5] Indeed, when the multipliers are large (positive or negative), the causal analysis of culture and institutions—for example, on economic activity—loses relevance to the study of the interactions between culture and institutions.

We specialize the abstract model of the joint evolution of culture and institutions into different examples of the interaction of elites and civil society (or of the class struggle) as in the classic sociology and political sciences, notably after Marx (1867), Pareto (1901), and Aron (1950a, 1950b). Indeed, these example models identify the distinct roles of culture, institutions, and their interactions in different societies of interests, elucidating several important themes in the literature and shedding light, as already noted above, on the role of culture in modulating the socioeconomic effects of democratization. More specifically, we illustrate

[5] See the *Handbook of Historical Economics* (Bisin and Federico 2021) for several surveys, especially Valencia (2021), Cantoni and Yuchtman (2021), Voth (2021), and Bisin and Moro (2021). See also Arroyo Abad and Maurer (2021) and Chen, Wang, and Yan (2022).

the explanatory power of the abstract model by studying the sustainability of extractive institutions and the formation of civic capital in two example societies.[6]

In the first society, elites are culturally heterogeneous, distinguished as *aristocracy* and *bourgeoisie*. The aristocrats cultivate their preferences for leisure while extracting resources from the workers via taxation. The bourgeois instead have the same preferences as the rest of civil society (workers), though more economic resources. In this society, we study the conditions under which the cultural and institutional dynamics maintain or reverse extractive institutions. We show that in such a society, the bourgeoisie might dominate the political process of the elites and have an interest in establishing less extractive institutions, that is, to allow for some form of political democratization. When it does, it chooses to devolve part of the fiscal authority to workers, indirectly committing institutions to a lower tax rate. This in turn induces workers to exert a higher labor effort, thereby spurring society's economic activity. On the other hand, depending on historical initial conditions, the democratization process might not have enough cultural and political space to be initiated, and society might remain locked in extractive institutions and limited economic activity. We show that in this society, culture and institutions are complements: the democratization process weakens the incentives of the aristocracy to transmit its own cultural trait (their preferences for leisure), and with a smaller aristocracy (a larger bourgeoisie), elites have higher incentives to devolve fiscal authority to workers.

In the second society, it is civil society that is culturally heterogeneous: only a fraction of its members is endowed with *civic capital*. Civic capital has beneficial effects on the functioning of public governance structures and hence on other members of civil society. Elites instead constitute a sort of *caste of bureaucrats*, exploiting opportunities for corruption from public good provision in society. In this society, we study conditions under which the cultural and institutional dynamics favors or hinders the accumulation of civic capital in society. We show that in this society, culture and institutions may act as substitutes: civic capital is more likely to spread when the degree of political representation in civil society is large and diffused, but the larger the diffusion of civic capital in society, the smaller the need to design institutional changes devolving formal power to prevent the misgovernance of public policies. An exogenous institutional change enlarging political representation—as a form of democratization—may end up having its effects mitigated by regressive dynamics of the accumulation of civic capital.

[6] Other themes this model has been specialized to study include the protection of property rights (Bisin and Verdier 2021), religious legitimacy (Bisin et al. 2021), cultural revivals (Iygun, Rubin, and Seror 2021), and the industrialization process (Touré, 2021).

The paper is structured as follows. Below we briefly discuss the related literature. In section II, we introduce separately the model of the dynamics of institutions and the model of the dynamics of culture. In section III, we study the joint dynamics of culture and institutions, and we introduce the cultural and institutional multipliers as tools for the empirical analysis. Section III ends with a discussion of three important extensions, allowing (i) for some forward-looking behavior in the institutional design process, with possibly slippery slope effects; (ii) for (strategic) actions/policies driving cultural dynamics, for example, the actions of cultural leaders; and (iii) for cultural heterogeneity in both political groups, possibly inducing a circulation of the elites, as in Mosca (1896) and Pareto (1916). In section IV, we introduce and study the two detailed example societies, illustrating some dynamical interactions of elites and civil society. Finally, in section V, we conclude.

Related literature.—We model institutions as a representation of the relative power of different political groups. This is in line with the pathbreaking series of contributions by Acemoglu, Johnson, Robinson, and others,[7] but it also diverges from it in several ways. In Acemoglu (2003) and Acemoglu and Robinson (2006), for example, institutions are a representation of political pressure groups exercising the power to control social choice, and institutional change takes the form of voluntary transfer of power across groups, typically under threat of social conflict. In this paper instead, we depart from the notion of political power as concentrated in one single political group. We represent institutions as Pareto weights associated with the different groups in the social choice problem. This allows us to view institutional change as more incremental (formally, a continuous rather than a discrete change in political control) than just revolutions and regime changes,[8] in line, for example, with the wealth of examples of institutional evolution through gradual and piecemeal changes in Mahoney and Thelen (2010).[9] It also allows us to enlarge the scope of our analysis from the study of transitions between autocracy and democracy and vice versa, which, for example, Acemoglu and Robinson (2000, 2001, 2006) concentrate on.

Our modeling of institutional change is also related to the cited contributions by Acemoglu, Johnson, and Robinson (2006) and others in that institutional change operates as a commitment mechanism (see also Jack and Lagunoff 2006). But our analysis focuses on mechanisms designed to

[7] See Acemoglu, Johnson, and Robinson (2006), and Acemoglu, Egorov, and Sonin (2021) for surveys. See also Przeworski (2004) for an early discussion and Bowles et al. (2021) and Levine and Modica (2021) for an alternative evolutionary approach based on external conflict.

[8] On the other hand, our analysis can be extended to account for (a smoothed formulation of) revolutions and regime changes (see Bisin and Verdier 2021).

[9] Relatedly, see Gradstein (2007, 2008) and Guimaraes and Sheedy (2016), who ground the study of institutions in the theory of coalition formation.

internalize inefficient political choices rather than to limit the threat of social conflict.[10] Indeed, our modeling of institutions is well aligned with North and Weingast's (1989) narrative about the historical events after the Glorious Revolution in seventeenth-century England, whereby institutions evolved to alleviate the Stuart monarchy's fiscal policy commitment problem. The process of institutional change we study in this paper is also characterized by some form of myopia to simplify the analysis.[11] Interestingly, however, myopic institutional change may be also factually motivated, for example, in the historical process that underlies the emergence of democracy (see Treisman 2017).[12]

As far as culture is concerned, we conceptualize it as preference traits, norms, and attitudes, and we allow for several social selection forces. In fact, the replicator dynamics we postulate can be microfounded from (i) evolutionary models using various payoff imitation protocols (Helbing 1992; Hofbauer 1995; Bjornerstedt and Weibull 1996; Weibull 1995); (ii) indirect evolutionary models of preference dynamics (Güth and Yaari 1992; Güth 1995; for applications to specific contexts: Alger and Weibull 2013; Besley 2017, 2020; Besley and Persson 2019, 2020); and (iii) evolutionary anthropology models of cultural transmission (Cavalli-Sforza and Feldman 1973, 1981; Boyd and Richerson 1985; for an economic approach with parental socialization choice: Bisin and Verdier 1998, 2000a, 2000b, 2001).[13] We should emphasize, however, that the notion of culture we adopt represents a relatively specific dimension of how culture is conceptualized in the social sciences. In cultural sociology, for example, culture is not thought of as being about values and preferences but rather, following Geertz (1973), about meaning. In this sense, culture is a tool kit of attributes agents draw on to accomplish and legitimize particular strategies of action (Swidler 1986; Alexander 2003; DiMaggio and Markus 2010).[14]

[10] For specific positions along these lines pertaining to the explanation of the extension of the franchise in early nineteenth-century England, see Acemoglu and Robinson (2000, 2001, 2006), Conley and Temimi (2001), and Lizzeri and Persico (2004).

[11] A fully forward-looking model of institutional change is analytically intractable when joined with cultural dynamics, though forward-looking institutional change per se is studied by Lagunoff (2009) and Acemoglu, Egorov, and Sonin (2015). In sec. IV, we extend our model to accommodate some forward-looking behavior to encompass slippery slope arguments.

[12] Specifically, Treisman (2017) argues that in the majority of the events he classifies, democracy has been the outcome of miscalculation and lack of anticipation of the effects of the process set in motion by institutional change. In several instances, the "incumbent initiates a partial reform . . . but cannot stop" (see table 2 in the paper), a representation that closely maps our modeling of myopic institutional change.

[13] See Bisin and Verdier (2011, 2021) for surveys and discussions and app. B for the formal derivations.

[14] Following this perspective, Acemoglu and Robinson (2021) represent culture as a hierarchical structure of a set of attributes and configurations reflecting specific associations between attributes. Depending on the nature and connectivity properties of the attributes, the cultural system is characterized by a certain degree of fluidity, namely, the span of

A number of papers study theoretically the implications of the interactions between culture and institutions for economic activities. These papers, however, typically each focus on a distinct context-specific instance of these interactions rather than on an abstract model of the political economy of elites and civil society as in this paper—notably, for example, work norms and the welfare state (Bisin and Verdier 2000b), norms of cooperation and legal systems (Tabellini 2008b), preference for patience and work ethics as well as labor markets in the Industrial Revolution (Doepke and Zilibotti 2008), trust and regulation (Aghion et al. 2010), organizational culture and incentives (Besley and Ghatak 2017; Besley and Persson 2020), civic culture and democratic institutions (Ticchi, Verdier, and Vindigni 2013; Besley and Persson 2019), and individualism and market organization (Davis and Williamson 2016).[15]

II. The Society

Consider a society with a continuum of agents separated into two political groups characterized by distinct economic resources, political power, and cultural traits. More specifically, for example, the groups represent *elites* and *civil society*, whereby elites have more resources or a different technology to obtain them.[16] Political groups are composed of agents with possibly heterogeneous cultural traits. For instance, members of the elites can be divided into aristocrats and bourgeois, or agents in civil society can be civic minded and non–civic minded (see sec. IV for specific examples along these lines). For analytical tractability, we assume that only one of the two groups is culturally heterogeneous,[17] and we consider dichotomous traits, indexed by $i = 1, 2$.

Let the choice of an agent in the culturally homogeneous group be denoted a and the choice of an agent in the culturally heterogeneous group with trait i be denoted a^i.[18] The government's socioeconomic policy choice is denoted p. Let actions and policy choices lie in compact real intervals. Let $\mathbf{a} = \{a, a^1, a^2\}$ denote the profile of actions and let $\mathbf{e} = (\mathbf{a}, p)$. Let λ be the fraction of the cultural heterogeneous political group in the

alternative configurations that can be generated through the system. Studying a population dynamics model of cultural transmission or diffusion where culture is richly defined, as in Acemoglu and Robinson (2021), is a challenging endeavor.

15 See also Lindbeck (1995), Bidner and Francois (2011), Alesina and Giuliano (2015), Benabou, Ticchi, and Vindigni (2015), Acemoglu and Robinson (2021), Bisin and Verdier (2021) Gorodnichenko and Roland (2021), and Persson and Tabellini (2021) for surveys.

16 Restricting the analysis to two groups avoids the issue of coalition formation in institutional design, a limitation of our analysis.

17 But see sec. III.C.3 for an extension relaxing this assumption.

18 At the level of abstraction of this section, we do not (need to) specify whether elites or civil society are culturally heterogeneous; we shall do it in the examples in sec. IV.

population, and let q denote the share of agents in this group with trait 1 (and $1 - q$ the share with trait 2).

Cultural traits are represented by preference traits, norms and/or conventions agents might abide to, ethnic and/or religious identities, and so on. The preferences of agents belonging to the homogeneous political group are represented by a utility function $u(\mathbf{a}, p, q)$, and the preferences of agents in the heterogeneous group with trait i are represented by $u^i(\mathbf{a}, p, q)$. Utility functions u and u^i should be interpreted as indirect utility functions, so that their dependence on q captures the effects of technologies and resources through the distribution of the population by cultural trait in the heterogeneous group. Their dependence on the whole profile of actions $\mathbf{a}$ captures the possible presence of externalities in the economy. A natural example of an externality operating through the distribution of cultural traits could be represented by preferences depending on the mean action in the population, $A = (1 - \lambda)a + \lambda(qa^1 + (1 - q)a^2)$. These utility functions can be denoted compactly as $u(\mathbf{e}, q)$ and $u^i(\mathbf{e}, q)$.

We conceptualize institutions as mechanisms through which social choices are delineated and implemented at equilibrium. Specifically, we model institutions as weights associated with the different political groups in the social choice problem that determines policy making at equilibrium. We denote with β the weight associated with the heterogeneous group (and with $1 - \beta$ the weight associated with the homogeneous group). In other words, political power β is distinct by socioeconomic group—elites and civil society—but independent of the distribution of cultural traits in the heterogeneous group. The welfare weights parameterize synthetically the relative structure of political power that exists between the relevant political groups, that is, their relative bargaining power in policy setting (the social choice problem), interpreted as a collective choice problem. This bargaining power across political groups can be related to population size but also, for example, to how autocratic are political institutions or—within democracies—to legislative processing, agenda setting rules, voting rights, geographic definition of electoral districts, and constitutional restrictions. In other words, we interpret the political process as a mechanism that potentially distorts the bargaining power of political groups with respect to population size. How much and in which direction this distortion is operated is endogenously determined in the model, as we shall see in section II.A. The relative political power associated with the different cultures inside the heterogeneous group is instead assumed to be represented by their relative share in the population, q for the culture with trait 1 and $1 - q$ for the culture with trait 2.[19]

[19] This formulation is chosen so as to maintain dichotomous political power weights, β.

In the collective choice problem, public policies are evaluated according to the following social welfare function, taking β as given:

$$W(\beta; \mathbf{a}, p, q) = (1-\beta)u(\mathbf{a}, p, q) + \beta[qu^1(\mathbf{a}, p, q) + (1-q)u^2(\mathbf{a}, p, q)]. \quad (1)$$

Importantly, we do not interpret the social choice problem normatively but rather as the indirect choice problem solved by the political process. Following our interpretation of welfare weights as the relative structure of political power in society, a natural interpretation of this problem is as an asymmetric Nash bargaining problem, where the asymmetry represents the relative bargaining power of the political groups, encoded by β; see Pearsall (1965) and Kaneko and Nakamura (1979) for microfoundations.[20]

A. *Societal Equilibrium and Welfare*

In this section, we introduce the concept of equilibrium for our society, given institutions β and given distribution by cultural traits in the heterogeneous group, q. We then study the welfare properties of equilibrium. We restrict the analysis to symmetric equilibria, where agents in the same group make the same choice.[21]

At an equilibrium in society, agents act noncooperatively with respect to each other and with respect to the economic policy choice. That is, given the policy choice p, $\mathbf{a}$ is a Nash equilibrium of the agents' choice problem. Economic policy is chosen to maximize the social welfare function, which encodes the relative power of the groups, without commitment: the policy maker cannot choose the policy p in advance of the choices of the economic agents. Formally, a *societal equilibrium* $\mathbf{e} = (\mathbf{a}, p)$ is a Nash equilibrium of the simultaneous game between agents and the policy maker in an institutional setup characterized by weights β and distribution by cultural trait q:[22]

$$\begin{aligned} p &\in \arg\max_p W(\beta; \mathbf{a}, p, q), \\ a &\in \arg\max_a u(\mathbf{a}, p, q), \qquad (2) \\ a^i &\in \arg\max_{a^i} u^i(\mathbf{a}, p, q), \ i = 1, 2. \end{aligned}$$

[20] This problem can also find appropriate microfoundations in probabilistic voting models: the outcome of two-party rent-seeking competition converges toward a policy that maximizes a weighted welfare function of the relevant groups, where the weights depend, e.g., on voting rights, group size, lobbying capacity, and so on (see Lindbeck and Weibull 1987; Dixit and Londregan 1996; Grossman and Helpman 1996, 2001; Persson and Tabellini 2000). In app. A, we formally develop microfoundations along these lines. Other microfoundations have been obtained in the political economy of trade literature (see Grossman and Helpman 1994).

[21] We cannot rule out other equilibria. Alternatively, the reader can think of an interpretation of the model with large agents.

[22] We do not explicitly denote equilibrium variables as such, e.g., with stars; the (small) costs in clarity are hopefully compensated by the avoidance of excessively cumbersome notation.

A societal equilibrium will generally not be efficient. Indeed, the agents' equilibrium choice $\mathbf{a}$ will not be efficient in general, given the policy choice p, because of the strategic interactions across agents and because of the externality directly embedded in the formulation of preferences. The policy choice p then adds a further layer of inefficiency because of lack of commitment.[23] More precisely, a formal welfare analysis of societal equilibrium requires defining the *societal optimum* $(\mathbf{a}^{\text{eff}}, p^{\text{eff}})$,

$$(\mathbf{a}^{\text{eff}}, p^{\text{eff}}) \in \arg\max \ W(\beta; \mathbf{a}, p, q), \tag{3}$$

and the equilibrium notion under commitment, the *societal commitment equilibrium*, $\mathbf{e}^{\text{com}} = (\mathbf{a}^{\text{com}}, p^{\text{com}})$, defined as the Stackelberg Nash equilibrium of the societal game,

$$\begin{gathered} p^{com} \in \arg\max \ W(\beta; \mathbf{a}, p, q), \\ \text{such that } a \in \arg\max_a \ u(\mathbf{a}, p, q), \\ a^i \in \arg\max_{a^i} \ u^i(\mathbf{a}, p, q), \ i = 1, 2. \end{gathered} \tag{4}$$

Making the dependence on (β, q) explicit, the societal equilibrium and the societal commitment equilibrium can be denoted, respectively, by

$$\mathbf{e}(\beta, q) = \begin{pmatrix} \mathbf{a}(\beta, q) \\ p(\beta, q) \end{pmatrix}$$

and

$$\mathbf{e}^{\text{com}}(\beta, q) = \begin{pmatrix} \mathbf{a}^{\text{com}}(\beta, q) \\ p^{\text{com}}(\beta, q) \end{pmatrix}.$$

For regularity, we assume that utility functions are such that $\mathbf{e}(\beta_t, q_t)$, $\mathbf{e}^{\text{com}}(\beta_t, q_t)$ are continuous functions.[24] It is straightforward then to show that the societal optimum, the societal equilibrium, and the societal commitment equilibrium are generally distinct and weakly ranked in terms of welfare.[25]

Proposition 1. Given β and q, the societal optimum weakly dominates the societal commitment equilibrium, which in turn weakly dominates the societal equilibrium.

This result is a straightforward consequence of the fact that for any (β, q), (i) problem (4), which defines a societal commitment equilibrium, is a constrained version of problem (3), which in turn defines a societal

[23] See also Acemoglu (2003) and Belloc and Bowles (2013, 2017) for models of inefficient institutional dynamics in a different context.

[24] See app. C for the obvious but stringent restrictions on fundamentals.

[25] We say that a couple $(\mathbf{a}, p)$ weakly (strictly) dominates another one if the first is weakly preferred by (strictly preferred by at least one of the) agents.

optimum; and (ii) any societal equilibrium satisfying (2) is always contained in the constrained feasible set of problem (4), which defines a societal commitment equilibrium.

To illustrate conditions determining whether the dominance relationships established in proposition 1 are weak or strict, it is convenient to consider a simple society facing redistributive policies, in which the role of externalities and of lack of commitment is clearly apparent. We introduce this society in figure 1 and use it as a running example in this section.

B. *Institutional Dynamics*

We postulate institutions evolving over time as a mechanism to alleviate the inefficiency that plagues the societal equilibrium, that is, the inefficiency due to the direct externality in the agents preferences and to the lack of commitment of the policy maker. In particular, while economic policies are chosen without commitment, society can commit to institutional change in the form of redistribution of political power across groups. Institutional change then occurs when redistributing power across groups leads to higher social welfare, evaluated with respect to the distribution of power prior to the change. The mechanism driving the institutional dynamics of society we postulate is therefore akin to optimal political delegation: more political power is delegated to the group that is better able (or has the highest incentives) to internalize the externalities affecting the policy game.[26]

More precisely and operationally, a given current set of institutions in period t, β_t,[27] induces a social preference order internalized by the policy maker at t and hence the policy choice $p(\beta_t, q_t)$ at equilibrium. But social welfare evaluated at weights β_t is highest under policy $p^{\mathrm{com}}(\beta_t, q_t)$, that is, under commitment. Future political institutions, β_{t+1}, are then designed at the end of period t to aim at a policy closer to $p^{\mathrm{com}}(\beta_t, q_{t+1})$. For a given society characterized by institutions β_t and an anticipated distribution by cultural trait q_{t+1}, we postulate dynamics of institutions driven by the function $P(\beta_t, q_{t+1}) \coloneqq p^{\mathrm{com}}(\beta_t, q_{t+1}) - p(\beta_t, q_{t+1})$, which is an indicator of the extent of the policy commitment problem that will be faced by such society at time $t+1$ under no institutional change. Specifically, the absolute value of $P(\beta_t, q_{t+1})$ indicates the intensity of the commitment problem, reflecting the distance between what can best be achieved under commitment and what is actually achieved at equilibrium. The sign of $P(\beta_t, q_{t+1})$, on the other hand, indicates the direction of institutional

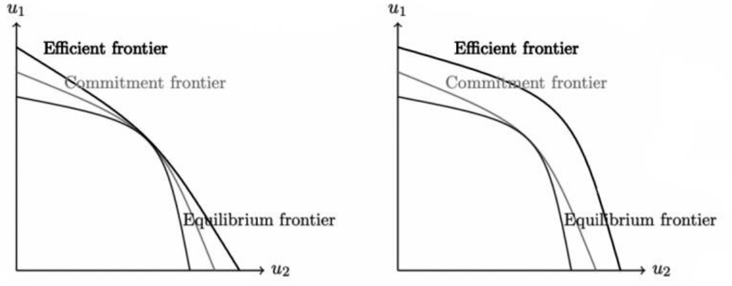
[26] This notion of political delegation has its conceptual roots in the analysis of incomplete contracts (in Grossman and Hart 1986), where ownership rights influence the efficiency of specific investments.

[27] We turn to an index t for an explicit notation for time.

Consider a society where policy is a purely redistributive fiscal policy between elites and workers and there is no cultural heterogeneity in either group. Assume taxes and subsidies are distortionary. The two groups are of the same size in the population. Each member of the elites (resp. each worker) is endowed with a production function h^e (resp. h) which maps his/her input a^e (resp. a) into output $h^e(a^e)$ (resp. $h(a)$). Input of elites is taxed linearly at rate p and the proceeds rebated to workers as a linear subsidy at rate r so as to satisfy the government budget constraint, $pa^e + ra = 0$. (Taxation of effort rather than output is unusual, but simpler in this context. The consumption of agents in the two groups s - that is, their output net of taxes and subsidies - is: $h^e(a^e) - pa^e$; $h(a) - ra$, with no externality; and $h^e(a^e) - pa^e$; $h(a) + v(a^e) - ra$, with externality $v(a^e)$.

A Utility Frontiers: No Externality. Unique β^* such that $\mathbf{e}(\beta^*) = \mathbf{e}^{eff}(\beta^*)$. Therefore, also $\mathbf{e}(\beta^*) = \mathbf{e}^{com}(\beta^*)$. For any $\beta \neq \beta^*$, $\mathbf{e}(\beta)$ is strictly dominated by $\mathbf{e}^{com}(\beta)$, which in turn is strictly dominated by $\mathbf{e}^{eff}(\beta)$.

B Utility Frontiers: Externality. $\mathbf{e}(\beta) \neq \mathbf{e}^{eff}(\beta)$, for any $\beta \in [0, 1]$. Unique β^E such that $\mathbf{e}(\beta) = \mathbf{e}^{com}(\beta)$. For any $\beta \neq \beta^E$, $\mathbf{e}(\beta)$ is strictly dominated by $\mathbf{e}^{com}(\beta)$, which in turn is strictly dominated by $\mathbf{e}^{eff}(\beta)$.



If the production choices of agents display no externalities - Quadrant a) - the *societal equilibrium* is efficient if and only if there is no fiscal policy intervention. More precisely, there exist a unique set of institutions β^* ($= 1/2$ if production technologies are symmetric) such that no redistributive policy is implemented at a *societal equilibrium* and the societal equilibrium is efficient. Any other set of institutions $\beta \neq \beta^*$ induces a redistributive fiscal policy, which is inefficient due to the distortionary effects of taxes and subsidies and to lack of commitment. Furthermore, for any $\beta \neq \beta^*$ the *societal equilibrium* is strictly dominated by the *commitment societal equilibrium* which is itself inefficient, but only because of the distortions. If, on the contrary, the production choice of elites agents has a positive externality on the income of workers - quadrant b), right - *societal equilibria* are always inefficient. There is a unique set of institutions β^E such that the *societal equilibrium* coincides with the *societal commitment equilibrium*. At β^E, the societal commitment equilibrium has the property that the marginal benefit of the externality is equal to the marginal cost induced by the distortionary subsidy. For any $\beta \neq \beta^E$ the *societal equilibrium* is strictly dominated by the *societal commitment equilibrium*, and both are inefficient.

FIG. 1.—Redistribution example: welfare analysis.

change in β_t needed to ameliorate the commitment problem. An abstract general dynamics for institutions can then be written as follows:

$$\beta_{t+1} - \beta_t = P(\beta_t, q_{t+1})\alpha(\beta_t, q_{t+1}), \tag{5}$$

where the dynamics $\beta_{t+1} - \beta_t$ is proportional to the extent of the commitment problem, $P(\beta_t, q_{t+1})$, and the proportionality factor, $\alpha(\beta_t, q_{t+1})$, is assumed to be strictly positive and, depending on β_t, q_{t+1} to capture different forms of nonlinearities.[28]

An interesting special case of these dynamics is useful to ground the conceptual structure supporting them. Consider future political institutions, β_{t+1}, designed at the end of period t to maximize the current social welfare function by means of future policy choices at $t + 1$. If we assume that institutional design is myopic—that is, institutions are designed for the future as if they would never be designed anew in the forward future—institutions at time $t + 1$ are designed at time t as a solution to

$$\max_{\beta_{t+1}} W(\beta_t; \mathbf{a}(\beta_{t+1}, q_{t+1}), p(\beta_{t+1}, q_{t+1}), q_{t+1}). \tag{6}$$

In words, the societal equilibrium induced by institutions β_{t+1} at $t + 1$ is chosen to maximize the social welfare induced by institutions β_t. The formal characterization of the resulting institutional dynamics has a fundamental structure in this case, with a clear and straight interpretation. We show in appendix A that in the solution of problem (6), β_{t+1} is chosen so that

$$p^{\text{com}}(\beta_t, q_{t+1}) = p(\beta_{t+1}, q_{t+1}), \tag{7}$$

unless equation (7) does not have a solution, in which case β_{t+1} is chosen at a corner, either 0 or 1. Note that, indeed, if we take a linear approximation (in continuous time), equation (7) is nested into (5), with $\alpha(\beta_t, q_{t+1}) = (\partial p(\beta_t, q_{t+1})/\partial \beta_t)^{-1}$.[29]

The properties of the dynamics of institutions β_t deriving from equation (5)—and (7)—that are most relevant in our subsequent analysis are collected in the following:[30]

Proposition 2. Given q, the dynamics of institutions in equation (5) have at least one stationary state. Any interior stationary state β^* obtains as a solution to $P(\beta, q) = 0$. The boundary stationary state $\beta = 1$ obtains when $P(1, q) > 0$, while the boundary stationary state $\beta = 0$ obtains when

[28] The dynamics of eq. [5] can more generally also account for (a continuous change version of) the institutional dynamics in Acemoglu and Robinson (2000, 2006), where delegation of power on the part of the elites does not serve to guarantee commitment policies or the internalization of externalities but rather the avoidance of social conflict; see Bisin and Verdier (2021) for a formal argument.

[29] In app. sec. C1, we provide straightforward but stringent sufficient conditions on fundamentals that guarantee the existence of a unique societal equilibrium. In such a case, the linearity of the policy objective function with respect to β_t implies that $p(\beta_t, q_t)$ is necessarily monotonic in β_t. Without loss of generality, we then take $p(\beta_t, q_t)$ increasing in β_t. This amounts to selecting the culturally heterogeneous group (with weight β_t) as the one preferring higher values of the policy p.

[30] A more complete global stability analysis is not particularly complex but tedious. We relegate it to app. B, proposition B1.

$P(0, q) < 0$.[31] In the continuous time limit, the dynamics satisfy the following properties:

- if $P(\beta, q) > 0$ for every $\beta \in [0, 1]$, then $\beta = 1$ is a globally stable stationary state;
- if $P(\beta, q) < 0$ for every $\beta \in [0, 1]$, then $\beta = 0$ is a globally stable stationary state;
- any boundary stationary state is always locally stable; and
- any interior stationary state β^* is locally stable if $\partial P(\beta^*, q)/\partial\beta < 0$.

We consider as an illustration the redistribution example introduced in figure 1, which we continue in figure 2.

The characterization of the stationary states of the dynamics of institutions we obtained in proposition 2 can be shown to imply that while the institutional dynamics do drive society toward efficiency, it is not generally the case that institutions are efficient in a stationary state or that institutions lead to Pareto improvements.[32] Indeed, the dynamics of institutions has an efficient stationary state iff there exist β^* such that $\mathbf{e}(\beta^*, q) = \mathbf{e}^{\text{eff}}(\beta^*, q)$. If no such β^* exists, all stationary states are inefficient. Furthermore, the dynamics of institutions does not generally converge to a stationary state that Pareto dominates the initial institutional state of society, β_0. More specifically, if the dynamics of β_t converges to an interior stationary state β, then $e(\beta, q) = e^{\text{com}}(\beta, q)$. But in this case, the societal commitment equilibrium utility frontier is (weakly) negatively sloped and the slope of the societal equilibrium frontier coincides with it at β. As a consequence, the dynamics from any β_0 in a neighborhood of the stationary state drives the utilities of the two groups in opposite directions. In the redistribution society with no externality, the dynamics converges to $e(\beta^*, q)$, which is efficient, but does not constitute a Pareto improvement for some initial institutional state β_0 close enough to β^*. In the case with the externality, the dynamics converges to $e(\beta^E, q)$, which is instead not efficient and, as in the case with no externality, does not constitute a Pareto improvement for some initial institutional state β_0 close enough to β^E.

C. *Cultural Dynamics*

We postulate a dynamics of the distribution of the population in the culturally heterogeneous group following a simple replicator dynamics. This functional form—a logistic equation—is the formal representation of

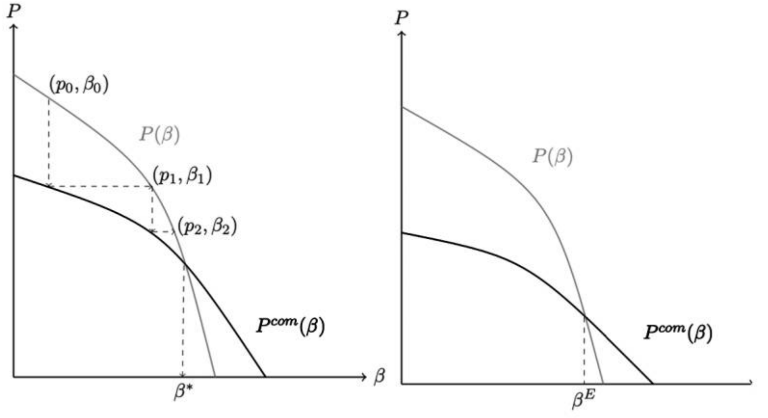
[31] Note that we arbitrarily define $\beta = 1$ ($\beta = 0$) as an *interior stationary state* if $P(\beta, q^i)|_{\beta=1} = 0$ ($P(\beta, q^i)|_{\beta=1} = 0$).

[32] For early analyses of institutions evolving toward efficiency, see Demsetz (1967) on property rights and Wittman (1989) on democracy.

Dynamics of Institutions: $\frac{\beta_{t+1}}{1-\beta_{t+1}} = \frac{\beta_t}{1-\beta_t} D(\mathbf{e})$, with $P(\beta) \gtreqless 0$ if $D(\mathbf{e}) \gtreqless 1$.

A Dynamics of Institutions: No Externality. At $\beta = \beta^*$, $P(\beta) = 0$, $D(\mathbf{e}) = 1$, $\epsilon^e + \epsilon = 0$; where $\epsilon^e = \mid \frac{p}{a^e} \frac{da^e}{dp} \mid_{p=p^{com}}$, $\epsilon = \mid \frac{r}{a} \frac{da}{dr} \mid_{r=r^{com}}$.

B Dynamics of Institutions: Externality. At $\beta = \beta^E$, $P(\beta) = 0$, $D(\mathbf{e}) = 1$.



In this society, in both the cases with and without the externality, Equation 7 governs the dynamics of institutions. In the case in which there's no externality in production, $P(\beta, q) \gtreqless 0$ if $\beta \gtreqless \beta^*$; while in the case with externality, this is the case if $\beta \gtreqless \beta^E$. In the first case, a policy of redistribution introduces distortions which reduce the level of the taxed input and stimulate the level of the subsidized input. Institutional change is then induced in such a way as to provide more power to the political group which is less distortionary at the margin. When $\beta_t < \beta^*$, the taxed group 1 is the group most affected by the redistributive distortions, in that the elasticity of this group input with respect to its tax, ϵ_1, is higher in absolute value than the elasticity of group 2 with respect to its subsidy, ϵ_2. Consequently more political decision rights are delegated to group 1, by increasing β_{t+1} with respect to β_t. Conversely when $\beta_t > \beta^*$, it is group 2 which is taxed and is most affected by the redistributive distortions. Consequently, Group 2 gets more political decision rights and β_{t+1} is decreased; see quadrant a), left. The case with externality is similar: increased political delegation is obtained by the group generating the positive externality as long as the marginal benefit of the externality is greater than the marginal cost induced by the distortionary subsidy. This is the case for $\beta < \beta^E$; see quadrant b), right. In the redistribution example, β_t converges to the unique interior stationary state, β^* and β^E, respectively, in the society without and with the externality.[a]

[a]See the Online Appendix for further details.

FIG. 2.—Redistribution example: dynamics.

several interesting distinct cultural selection processes, as we noted when discussing the related literature in section I. Formally, given β_{t+1}, the dynamics of the distribution by cultural trait q_t are governed by a difference equation of the following form:

$$q_{t+1} - q_t = q_t(1 - q_t)S(\beta_{t+1}, q_{t+1}), \tag{8}$$

where $S(\beta_{t+1}, q_{t+1})$ represents the relative strength of trait $i = 1$ in terms of its ability to spread in the population. Reflecting an abstract social selection process on cultural traits or norms of behaviors, $S(\beta, q)$ depends on the societal equilibrium set of actions and policy $\mathbf{e}(\beta, q) = [\mathbf{a}(\beta, q); p(\beta, q)]$. Typically, $S(\beta, q)$ takes the form

$$S(\beta, q) = h^1(\mathbf{e}(\beta, q), q) - h^2(\mathbf{e}(\beta, q), q), \tag{9}$$

where $h^i(\mathbf{e}, q)$ is an appropriate *cultural fitness* function of trait i in the population. In a pairwise comparison random matching imitation context (Weibull 1995), for instance, $h^i(\mathbf{e}, q)$ is simply proportional to the utility $u^i(\mathbf{e}, q)$ of agents of type i. In an indirect evolutionary approach, $h^i(\mathbf{e}, q)$ represents the material fitness at the societal equilibrium $\mathbf{e}(\beta, q)$ for agents of type i, that is, with preferences $u^i(\mathbf{e}, q)$. In the cultural transmission models by Bisin and Verdier (2000a, 2000b, 2001), paternalistic parents of the two cultural types spend costly resources to bias the process of preference acquisition of their children. In this case, unpacking the notation of utility functions so that $u^i(\mathbf{a}, p, q) = u^i(a^i, a^j, p, q)$ and denoting $\Delta V^i(\beta, q) = u^i(a^i, a^j, p, q) - u^i(a^j, a^j, p, q)$, one gets

$$h^1(\mathbf{e}, q) = w((1 - q)\Delta V^1(\beta, q)) \text{ and } h^2(\mathbf{e}, q) = w(q\Delta V^2(\beta, q)), \tag{10}$$

where $w(\cdot)$ is an increasing function. In words, the cultural fitness of trait i is increasing in the socialization gain of parents of trait i.

We assume for regularity that $S(\beta, q)$ is a continuous function. Generally, and independent of the underlying specific cultural selection process, we can then characterize the dynamics of the distribution by cultural trait q_t as follows.[33]

Proposition 3. The dynamics of the distribution of culture q_t have at least the two boundaries as stationary states, $q = 0$ and $q = 1$. Any interior stationary state $0 < q^* < 1$ obtains as a solution to $S(\beta, q) = 0$, In the continuous time limit, the dynamics satisfy the following properties:

- if $S(\beta, q) > 0$ for every $q \in [0, 1]$, then q_t converges to $q = 1$ from any initial condition $q_0 > 0$;
- if $S(\beta, q) < 0$ for every $q \in [0, 1]$, then q_t converges to $q = 0$ from any initial condition $q_0 < 1$;
- if $S(\beta, 1) > 0$, then $q = 1$ is locally stable;
- if $S(\beta, 0) < 0$, then $q = 0$ is locally stable; and
- any interior stationary state $q^*(\beta)$ is locally stable if $\partial S(\beta, q^*)/\partial q < 0$.

[33] We collect here the properties of the dynamics of culture that are most relevant in our subsequent analysis. We relegate a more complete analysis to app. B, proposition B2.

When $S(\beta, q)$ has the form in (9), any interior steady state $q^*(\beta)$ satisfies $h^1(\mathbf{e}(\beta, q^*), q^*) = h^2(\mathbf{e}(\beta, q^*), q^*)$. Furthermore, when (10) is satisfied, any interior cultural stationary state $q^*(\beta)$ is obtained as a solution to[34]

$$\frac{\Delta V^1(\beta, q)}{\Delta V^2(\beta, q)} = \frac{q}{1 - q}. \tag{11}$$

III. Joint Evolution of Culture and Institutions

In this section, we study the dynamics of culture and institutions in the society introduced in section II, highlighting conditions under which interesting qualitative dynamical paths—such as dependence on initial conditions, limit cycles, or other oscillatory dynamics—may or may not arise. We then introduce the cultural and institutional multipliers as useful tools for the analysis of these dynamics. We draw implications for the study of the effects of culture and institutions on economic variables of interest as a complementary tool to the causal methods largely adopted in historical economics and particularly in persistence studies. Finally, we discuss some relevant extensions.

A. Dynamics

The model of cultural and institutional change introduced in section II delivers dynamics governed by the system of difference equations:

$$\beta_{t+1} - \beta_t = P(\beta_t, q_{t+1})\alpha(\beta_t, q_{t+1}), \tag{5}$$

$$q_{t+1} - q_t = q_t(1 - q_t)S(\beta_{t+1}, q_{t+1}). \tag{8}$$

Recall that $\alpha(\beta_t, q_{t+1}) > 0$. Any interior stationary state of the system ((5), (8)), (β^*, q^*), solves

$$P(\beta, q) = S(\beta, q) = 0. \tag{12}$$

While the explanatory power of this system is best manifested once it is specialized to the study of phase diagrams in specific societies,[35] in section IV, a series of results can be obtained even at this level of generality.

With regard to stationary states, while it is not always the case that an interior stationary state exists, we can show the following:

PROPOSITION 4. The dynamical system ((5), (8)) has at least one stationary state.

[34] In app. sec. B4.3, we show that for every value of β, the cultural dynamics are monotonic and converge toward a unique interior stationary state $q^*(\beta)$ under strong enough cultural substitution (Bisin and Verdier 2001), i.e., when members of the cultural minority have higher marginal incentives to engage in socialization than members of the majority.

[35] See Bisin and Verdier (2021) for a discussion of the role of phase diagrams in historical economics and several examples.

Clearly, multiple stationary states are possible in the general nonlinear system. In section IV, we will study interesting examples of nonergodic dynamics, where different stationary states are reached by different basins of attractions in the space of initial conditions.

With regard to local stability, we take a continuous time approximation. In this case, the formal characterization of the (possibly complex) dynamics of a two-dimensional system are of course well understood. We concentrate on identifying conditions that have clear interesting interpretations in terms of the properties of cultural and institutional change. To this end, we assume for regularity the following separability condition: $u^i(\mathbf{a}, p, q) = v^i(a^i, p) + H^i(\mathbf{a}, p, q)$ in the following analysis.

First of all, we have the following:

Proposition 5. The condition

$$\frac{\partial P(\beta^*, q^*)}{\partial \beta}, \frac{\partial S(\beta^*, q^*)}{\partial q} < 0 \tag{13}$$

is sufficient to guarantee no limit cycles in the neighborhood of an interior stationary state (β^*, q^*) of the dynamical system ((5), (8)).

Condition (13) is an implication of the Bendixon negative criterion and has a clear intuitive interpretation. It requires that in a neighborhood of the stationary state, institutional change has social decreasing returns as a mechanism to internalize the externalities of the equilibrium, $\partial P(\beta^*, q^*)/\partial \beta < 0$, and similarly that the selective forces driving cultural change also have decreasing returns, $\partial S(\beta^*, q^*)/\partial \beta < 0$, hence favoring cultural diversity in a neighborhood of q^*.[36] General conditions for the existence of limit cycles are not instructive at this level of generality, but we discuss interesting robust examples with a predator-prey interpretation in section III.C.3.

Condition (13) guarantees a substantial reduction in the complexity of the dynamics of the system but is not sufficient for local stability, which instead also requires conditions involving the complementarity or substitutability of culture and institutions in a precise sense that we define next.[37] Intuitively, we say that culture and institutions are complements (substitutes) if an increase in the relative political power of the culturally heterogeneous group β leads to an increase in the relative size of the population with one of its cultural traits—say, q, the relative size of the population with trait 1—which in turn leads to an increase (decrease) in the political power of the culturally heterogeneous group. Formally, let $\beta = \beta(q)$ be the stationary state manifold associated with the equation

[36] This is the case, e.g., under cultural substitution in cultural transmission models (see n. 34).

[37] The extension of this analysis to corner stationary states is notationally involved in the general case, with minor additional insights.

in (5), that is, such that $P(\beta(q), q) = 0$. Let $q(\beta)$ be the stationary state manifold associated with equation (8), that is, such that $S(\beta, q(\beta)) = 0$. Clearly, any interior stationary state lies at the intersection of the manifolds $\beta(q)$ and $q(\beta)$. We then say the following:

Institutional and cultural dynamics are locally complementary at an interior stationary state (β^*, q^*) when the stationary state manifolds $\beta(q)$ and $q(\beta)$ of the dynamical system ((5), (8)) have slopes of the same sign at (β^*, q^*):

$$\operatorname{sign}\left(\frac{d\beta(q^*)}{dq}\right) = \operatorname{sign}\left(\frac{dq(\beta^*)}{d\beta}\right). \tag{14}$$

Conversely, they are locally substitutes at (β^*, q^*) when the slopes have opposite signs.[38]

Indeed, $d\beta(q)/dq$ represents the relationship between the size of the population with trait 1 and the ability of the culturally heterogeneous group to internalize the externality at equilibrium when social welfare favors the group's policy preferences, while $dq(\beta)/d\beta$ represents the relationship between the culturally heterogeneous group's weight in social welfare and the strength of cultural diffusion of trait 1. When, by way of illustration, these two terms are both positive at a stationary state (the negative case is symmetric), institutional and cultural dynamics are locally complementary in the sense that the political power of the heterogeneous group breeds into the size of trait 1, which in turn breeds into political power of the heterogeneous group in a dynamic feedback.[39] We now show how complementarity and substitutability turn out to be determinant factors governing the dynamics of culture and institutions and the local stability of dynamical system ((5), (8)).

Proposition 6. Suppose that condition 13 is satisfied. Then local substitutability of the institutional and cultural dynamics at an interior steady state (β^*, q^*) of the dynamical system ((5), (8)) is sufficient for local stability. Under local complementarity, instead, local stability obtains if

$$\frac{d\beta(q^*)}{dq}\frac{dq(\beta^*)}{d\beta} < 1. \tag{15}$$

Furthermore, in this case, the local dynamics show no converging oscillatory dynamics.

When both institutional and cultural change display decreasing returns at the stationary state—when condition (13) is satisfied—substitutability

[38] All these conditions are useful in applications, as we will show in sec. IV, in that they are generally easy to check.

[39] When interior cultural stationary states $q^*(\beta)$ are obtained as a solution to (11)—i.e., by cultural transmission, as in Bisin and Verdier (2000a, 2000b, 2001)—the complementarity condition on the slopes of $\beta(q)$ and $q(\beta)$ can be shown to require that $d(\Delta V^1(p)/\Delta V^2(p))/dp$ and $\partial P(\beta^*, q^*)/\partial q$ have the same sign (see app. B for details).

implies that culture and institutions dampen each other on a convergence path, thereby guaranteeing local stability, though possibly inducing oscillatory dynamics. Under complementarity, instead, culture and institutions reinforce each other's dynamics. A strong complementarity might therefore have amplification effects, inducing dynamics that diverge toward the corners of the dynamical system: condition (15) is effectively a bound on the strength of the complementarity of culture and institutions, which is required for stability. The (monotonicity of) reinforcement effects of complementary institutional and cultural dynamics, however, rule out oscillatory dynamics along a convergence path.

B. *The Cultural and Institutional Multipliers*

In this section, we introduce the concepts of cultural (institutional) multiplier, which we will then exploit when we specialize our analysis to the study of phase diagrams in specific societies in section IV. We define the *cultural* (*institutional*) *multiplier* as the ratio of the long-run change in institutions (culture) relative to the counterfactual long-run change that would have happened had the cultural composition (institutional setup) of society remained fixed. For ease of exposition, we shall concentrate on the cultural multiplier under the understanding that symmetric arguments and conditions hold for the institutional multiplier. In fact, an implication of the following analysis is that the cultural and the institutional multiplier have the same sign, which depends on the complementarity/substitutability of culture and institutions.

Consider a long-run socioeconomic variable that depends indirectly (in reduced form) on culture and institutions, $A(\beta, q)$. In fact, to illustrate the analysis, we consider the case in which both β and q have a positive effect on A.[40] We distinguish in turn two comparative dynamics exercises: the first one where a change $d\gamma$ in a parameter perturbs locally an interior stable stationary state (β^*, q^*), inducing a dynamical convergence path in culture and institutions and hence in A; and the second one where we follow the dynamics of culture and institutions from an initial condition (β_0, q_0) in the basin of attraction of a stationary state (β^*, q^*) and hence from $A(\beta_0, q_0)$ to $A(\beta^*, q^*)$.

Local change $d\gamma$.—Adding explicit reference to γ in the notation, we normalize the arbitrary components of the comparative dynamics environment we study so that a positive change in γ induces a process of convergence to a new steady state characterized by a larger societal equilibrium policy p (through a larger β).[41] In this context, the change

[40] See app. B for the complete analysis, when A depends on agents' actions $\mathbf{a}$ and government policies p.

[41] Formally, we sign the effects of a $d\gamma > 0$ so that it increases, locally at the steady state, both the policy p as well as the extent of the social externality or commitment problem: $dp^{\mathrm{com}}(\beta^*, q^*; \gamma)/d\gamma > dp(\beta^*, q^*; \gamma)/d\gamma > 0$. Also, without loss of generality, we let members

in the long-run variable of interest A has two interrelated components $dA/d\gamma = (\partial A/\partial\beta)(d\beta/d\gamma) + (\partial A/\partial q)(dq/d\gamma)$, where $d\beta/d\gamma = (\partial\beta/\partial\gamma) + (\partial\beta/\partial q)(dq/d\gamma)$ and $dq/d\gamma = (\partial q/\partial\gamma) + (\partial q/\partial\beta)(d\beta/d\gamma)$ are the institutional and cultural components, respectively. Then:

The cultural multiplier on institutional change at a locally stable interior steady state (β^*, q^*), m_{SS}, is

$$m_{SS} = \left(\frac{d\beta^*}{d\gamma}\right) \Big/ \left(\frac{\partial\beta^*}{\partial\gamma}\right)_{q=q^*} - 1 = \frac{\partial\beta}{\partial q}\frac{dq}{d\gamma}. \tag{16}$$

It follows that whether the multiplier is indeed positive crucially depends on culture and institutions being complements or substitutes. We can then show the following:

Proposition 7. Local complementarity (substitutability) of the institutional and cultural dynamics at an interior stable stationary state (β^*, q^*) is sufficient for the cultural multiplier m_{SS} at (β^*, q^*) to be positive (negative). The resulting cultural multiplier on the long-run variable A is $(\partial A/\partial q)(dq/d\gamma)$, which is positive when m_{SS} is positive. In the complementarity case, under our normalization, an increase in γ is set to induce an increase in β that, because of complementarity, is reinforced by an increase in q, inducing positive feedback dynamics. Any exogenous institutional change, through an increase in γ, is amplified by the associated cultural dynamics that interact with institutions, leading to $m_{SS} > 0$. Conversely, institutional changes would be hindered by cultural changes (i.e., the cultural multiplier m_{SS} is negative) when culture and institution are substitutes, that is, when the slopes of $\beta(q)$ and $q(p)$ have opposite signs.

Global dynamics from (β_0, q_0).—Consider an initial condition (β_0, q_0) in the basin of attraction of a stationary state (β^*, q^*). In this case, the full dynamics of culture and institutions from (β_0, q_0) converges by construction to (β^*, q^*); that is, in particular, institutions converge to $\beta^* = \beta(q^*)$ and the long-run variable to $A(\beta^*, q^*)$. In the counterfactual case in which the cultural composition of society had remained fixed, the dynamics of institutions would have converged to $\beta(q_0)$ and A to $A(\beta(q_0), q_0)$. In such a case, we say the following:

The cultural multiplier on institutional change m_{DD} from initial condition (β_0, q_0) in the basin of attraction of a stationary state (β^*, q^*) is

$$m_{DD} = \frac{\beta(q^*)}{\beta(q_0)} - 1. \tag{17}$$

The formal analysis of m_{DD} requires distinguishing between the local complementarity just defined and a more stringent global complementarity, defined as follows:

---

of the heterogeneous political group (with institutional power β) aim at a relatively larger policy level, p: $\partial p(\beta^*, q^*; \gamma)/\partial\beta > 0$.

Institutional and cultural dynamics are globally complementary (substitutes) when the steady state manifolds $\beta(q)$ and $q(\beta)$ have slopes of the same sign (opposite signs) for all values $(\beta, q) \in [0, 1]^2$.
The next result shows that even in the context of global analysis, whether the multiplier is indeed positive crucially depends on culture and institutions being complements or substitutes.

PROPOSITION 8. Under global complementarity of the institutional and cultural dynamics, the cultural multiplier m_{DD} from initial condition (β_0, q_0) in the basin of attraction of (β^*, q^*) has the same sign as $(q^* - q_0) \cdot (d\beta(q)/dq)$.

The resulting cultural multiplier on the long-run variable A is $(A(\beta(q^*, q^*))/(A(\beta(q_0, q_0)) - 1$, which also is positive when m_{DD} is positive. As an illustration, suppose that culture and institutions are complements in the sense that $d\beta(q)/dq$ and $dq(\beta)/d\beta > 0$. Consider the process of convergence along a transition path from (β_0, q_0) to (β^*, q^*) with, say, $q_0 < q^*$. Because of global complementarity between institutions and culture, one cannot have dampened oscillations in the basin of attraction of (β^*, q^*), and q_t increases monotonically from q_0 toward q^*. Along that transition path, this involves changes in β and q that reinforce each other: an increase in β induces a further increase in q, which in turn feedbacks positively on the institutional weight β. Upon convergence, then $\beta^* > \beta(q_0)$, where $\beta(q_0)$ is the counterfactual institutional steady state with culture fixed at q_0. This process therefore implies a positive institutional (and cultural) multiplier m_{DD} from the initial condition (β_0, q_0).

The cultural multipliers m_{DD} and m_{SS} illustrate different perspectives of the interactions between institutions and culture. The cultural multiplier m_{DD} encapsulates properties along the transition path of the joint dynamics between institutions and culture for a given society over time. Conversely, the cultural multiplier m_{SS} exhibits the steady state interactive effects between institutions and culture related to exogenous variations of some parameter γ. It therefore emphasizes the joint long-run effects of culture and institutions that may be observed across societies. In both cases, a positive cultural multiplier indicates the reinforcing effect of a change in institutions at a societal equilibrium due to the endogeneity of culture.[42]

C. *Extensions*

In this section, we discuss three important extensions of our model, allowing for (i) forward-looking behavior in the institutional design process,

[42] Another interesting class of comparative dynamics exercises consists of a change in a parameter γ, which induces a change in the basin of attraction of two distinct stationary states of the system. While general results are difficult to obtain in these cases, in sec. V we discuss the empirical implications of this class of exercises for the casual analysis of average treatment effects (see also Bisin and Moro 2021).

(ii) strategic cultural actions/policies, and (iii) cultural heterogeneity in both political groups. These extensions give rise to interesting phenomena, such as, respectively, (i) slippery slopes slowing down the pace of institutional change, (ii) cultural leaders driving cultural dynamics in the population, and (iii) predator-prey cycles representing a circulation of the elites.

1. Forward-Looking Institutional Change: Slippery Slope Effects

Institutional design is myopic in our model; that is, institutional change at time t from β_t to β_{t+1} is predicated under the assumption that institutions β_{t+1} will not change in the future. In other words, the mechanism driving institutional change does not anticipate the actual dynamics represented by equation (5) from β_{t+1} to β_{t+2}, β_{t+3}, and so on. In particular, it could very well be that with respect to the social welfare order associated with institutions β_t, these dynamics will lead to dominated outcomes. An institutional change mechanism with better forward-looking ability might prevent or mitigate the logic of this institutional slippery slope, slowing down the pace of change or even stopping it altogether.

Our model can be extended in this direction to account for some forward-looking institutional change. Fixing the cultural profile q prevailing in society, we consider as an illustration a one-step forward-looking behavior: given institutions β_t in period t, institutions β_{t+1} are chosen to maximize the implied social welfare ordering at t, anticipating institutions β_{t+2} induced by equation (5).[43] Under forward-looking institutional change, institutions β_t balance the policy commitment gains from strategic delegation to β_{t+1} against its costs, but differently from the myopic case, it also takes into account the benefits and costs of delegation from β_{t+1} to β_{t+2}. In this context, with strong enough regularity conditions to guarantee convexity of the institutional change problem, we show in appendix A that an interior stable institutional steady state β^* under the myopic institutional dynamics in (the continuous time limit of) equation (5) is also a steady state of the one-step forward-looking institutional dynamics. More interestingly, the one-step forward-looking institutional dynamics change converges to that steady state β^* but at a reduced speed to mitigate the costs of the slippery slope in the dynamics of institutions. When the convexity conditions are not satisfied, on the other hand, institutional change might not necessarily imply marginal institutional adjustments, and the forward-looking dynamics may get stuck at a point that remains far away from the myopic institutional steady state β^*.[44]

[43] Of course, K-step forward mechanisms can be characterized similarly.

[44] See Acemoglu, Egorov, and Sonin (2015) for a related analysis but in a context characterized by a discrete—hence not convex—set of possible institutions.

2. Strategic Cultural Policies: Cultural Leaders

In our model, cultural dynamics are the results of evolutionary selective forces emerging from fully decentralized mechanisms. This perspective does not take into account the fact that cultural change can itself be influenced by centralized public institutions, such as states, churches, clans, and community leaders. A recent literature has considered the role of these centralized cultural transmission agents in different settings.[45] An extension of our analysis along these lines would allow for the choice of some form of group-level socialization effort, partly internalizing the effect of such effort on the dynamics of cultural change. In such a case, the joint dynamics of culture and institutions may result as the outcome of a dynamic game between public policy institutions, determining strategically the evolution of institutions β_t, and group-level cultural institutions, determining instead the evolution of cultural traits q_t.

While a full extension along these lines is analytically intractable, insights from the aforementioned literature on cultural leaders and comparative statics in games suggest two implications. First, the joint evolution of culture and institutions will have important forward-looking dimensions along both the institutional and the cultural dynamics, creating important sources of multiple dynamic paths and nonergodicity. Second, one may also expect the existence of additional positive (negative) cultural and institutional multiplier effects when culture and institutions are strategic complements (substitutes).

3. Cultural Heterogeneity in Both Groups: Circulation of the Elites

In section III.A—proposition 5—we have derived conditions guaranteeing no limit cycles in the neighborhood of an interior steady state. On the other hand, the existence of cycles is of general interest in our context. Pareto (1916) and Mosca (1896) before him have notably argued that the cycles in the distribution of political power across different groups are a structural property of societal dynamics, a property they refereed to as *circulation of the elites*. In this section, we construct conditions under which such cycles can be obtained in our model. To this end, we extend our analysis to study a society where both political groups, elites and civil society, are culturally heterogeneous. Interesting, the cycles we obtain have a predator-prey interpretation (Lotka 1920, 1925; Volterra 1926); that is, the power and size of one elite group happens at the expenses of the other. We shall illustrate the model next, referring to appendix A for the formal details.

[45] See, e.g., Acemoglu and Jackson (2015), Hauk and Mueller (2015), Verdier and Zenou (2015, 2018), Carvalho (2016), Prummer and Siedlarek (2017), Almagro and Andres-Cerezo (2020), and Carvalho and Sacks (2021).

More precisely, consider two culturally heterogeneous political groups denoted E and V of respective size λ and $1 - \lambda$. More precisely, in political group E, there are two cultural subgroups denoted E_1 and E_2 in proportions q_E and $1 - q_E$. Similarly in political group V, there are two cultural sub-groups V_1 and V_2 in proportions q_V and $1 - q_V$. Let β encode the political power of group E. Suppose that institutional change is (infinitely) faster than cultural change, so that β is always on the steady state manifold $\beta = (q_E, q_V)$ for all $(q_E, q_V) \in [0, 1]^2$.

We show in appendix A that the key assumption to obtain predator-prey dynamics is that culture and institutions are cultural complements with respect to group E and substitutes with respect to group V: that is, $\beta(q_E, q_V)$ is increasing in q_E and decreasing in q_V.[46] In other words, an increase in the size of the prey cultural subgroup (group E_1 in E), induces a (more salient) externality to be internalized, with institutional change increasing the political power of group E, β; on the contrary, an increase in the size of the predator cultural subgroup (group V_1 in V), q_V, induces an externality to be internalized decreasing β. Under these conditions, it is easy to see that an increase in q_E (the prey) induces an increase in p and hence in q_V (the predator), but the increase in q_V (the predator) induces a decrease in p, which in turn feeds on the fraction of the prey q_E. This is the predator-prey mechanism that admits sustained oscillations and limit cycles under several well-known conditions.

IV. Political Economy of Elites and Civil Society

In this section, we work out two specific model societies, simple but rich enough to display some interesting cultural and institutional dynamics within the political economy of elites and civil society.[47] More specifically, we intend these two examples as illustrations of the explanatory power of the abstract model, focusing in turn on the transition from extractive to inclusive institutions (democratization) and on the formation of civic capital.[48]

We interpret the analysis in these examples as shedding some light on the role of culture in mediating democratization processes and on the historical conditions determining the sustainability of democratization processes leading to economic growth. In this sense, the interactions between culture and institutions in these model societies speak to the

[46] Appropriately extending the definitions in sec. III, we find that these conditions require that $\partial P(\beta, q_E, q_V)/\partial q_E > 0$ and $\partial P(\beta, q_E, q_V)/\partial q_V < 0$.

[47] In these examples, for concreteness, we adopt the institutional dynamics derived in eq. (7) and the cultural dynamics in eq. (10). Also, we impose and exploit various regularity conditions without explicit mentioning them. We refer to the online appendix for all details.

[48] Relatedly, Bisin and Verdier (2021) specialize the abstract model in this paper to study the protection of property rights and Bisin et al. (2021) to study religious legitimacy. See also Iygun, Rubin, and Seror (2021) on cultural revivals and Touré (2021) on the industrialization process.

puzzling empirical results on the relationship between democracy and economic activity (Acemoglu et al. 2008), showing how the causal effects of economic activity—at the bottom of Lipset modernization theory (Lipset 1959)—disappear once historical factors influencing both income and democracy are accounted for.[49]

A. *Elites, Workers, and Extractive Institutions*

In this section, we study a society where the power of a political group—specifically, the elites—is exercised by extracting resources from the other group—the civil society, represented by workers—for example, via taxation. We study in particular conditions under which the cultural and institutional dynamics in this society maintain their extractive character or else reverse to more inclusive forms. Extractive institutions have been shown—in the extensive work spurred by Acemoglu, Johnson, and Robinson (2001) and Acemoglu and Robinson (2001, 2010)—to represent one of the major obstacles that has hindered economic growth and ultimately prosperity across history.[50] At the same time, the reversal to more inclusive institutional forms can be interpreted as a process of democratization. In this sense, we study how, depending on historical conditions, a democratization process away from an autocratic regime may or may not turn out not to be sustainable and lead to socioeconomic prosperity.[51]

Consider a society where members of the elites and workers have different resources and different cultural traits. Specifically, with regard to resources, elites are endowed with an initial (prior to production) endowment $s > 0$. Workers are not, and hence they can consume only off of their production. Furthermore, workers face a survival constraint, a minimum level of consumption necessary for survival. With regard to culture, the elites are culturally heterogeneous, distinguished as aristocracy and bourgeois. In particular, the preferences of the members of the aristocracy

[49] But see Benhabib, Corvalan, and Spiegel (2013).

[50] See, e.g., Engerman and Sokoloff (2002), Lange, Mahoney, and vom Hau (2006), Acemoglu and Robinson (2010), and Ogilvie (2021) for discussions of extractive and inclusive institutions in economic history. Ogilvie and Carus (2014) provide a good survey of the literature. For analyses on specific regions, see, e.g., Nunn (2007, 2008), Huillery (2009), Nunn and Wantchekon (2011), Michalopoulos and Papaioannou (2013, 2016), and Tadei (2018) for Africa; Dell (2010), Arroyo Abad and van Zanden (2016), and Paredes-Fuentes (2016) for Latin America; Carvalho and Dippel (2016) for the Caribbean Islands; Iyer (2010), and Bogart and Chaudhary (2019) for South Asia; and Dell and Olken (2020) for Indonesia.

In particular, while very stylized, this example identifies a fundamental role of the bourgeoisie in this process, foreshadowing a representation of either the formation of inclusive institutions protecting property rights in England (McCloskey 2006, 2010, 2017) or the maintenance of slavery institutions in the Roman Empire (Schiavone 1996 [2020]).

[51] We use this terminology somewhat loosely, referring to autocratic and democratic societies even though we do not model directly institutional forms but rather only the relative power of elites and workers in the political economy process.

are shaped by cultural norms that let them value leisure greatly, more than bourgeois. Workers have instead the same preferences as the bourgeois. Hence in equilibrium, the bourgeois and the workers will work, and the aristocrats will generally eschew labor and constitute a *leisure class*. In this society, taxes on labor income are purely extractive, being redistributed per capita to the members of the elites.[52] Furthermore, since members of the aristocracy do not work, (fiscal) institutions are extractive in that only workers bear the weight of fiscal policy.[53]

Institutions lack commitment; that is, fiscal authorities choose the tax rate without internalizing its effect on labor effort. This gives institutions generally an incentive to tax labor excessively. As a consequence, in this society, the elites—the bourgeoisie, especially—might have an interest in establishing less extractive institutions by delegating part of the fiscal authority to workers, that is, they might have an interest in the process of democratization of society. This would indirectly commit institutions to a lower tax rate, in turn inducing workers to exert a higher labor effort, contributing to higher income and public good production. Delegating fiscal authority to the workers (who have preferences closer to those of the bourgeois), however, weakens the incentives of the aristocracy to transmit its own culture and hence reduces the size of the leisure class; in turn, a smaller leisure class augments the incentives of the elites to delegate fiscal authority to workers. This form of complementarity drives the equilibrium dynamics of culture and institutions in this society.

Depending on their distribution by cultural trait, when the elites exert a dominant political control on the fiscal authority in society, they might impose a tax rate such that workers are constrained to subsistence, leading society into an autocratic extractive regime.[54] The institutional dynamics of this economy will in general be *nonergodic*, depending crucially on initial conditions. Only when the initial institutional setup guarantees enough control on fiscal authority on the part of the workers will the institutional dynamics tend to democratization, transitioning away from the extractive regime. Interestingly, this transition will generally induce the formation of a sizable bourgeoisie. As well, it is also the case that a larger bourgeoisie at the initial conditions favors the transition away from the extractive regime.

The detailed analysis of this society follows. Accordingly with the notation in section II, members of the elites (the culturally heterogeneous

[52] This is not substantial to the analysis. It is straightforward to allow taxes to finance a public good consumption valued by both groups.

[53] We assume that the bourgeois, as members of the elites, are not taxed; if they were, the mechanisms for the transition away from extractive institutions we focus on in this example society would be even stronger.

[54] The survival constraint can be binding only for workers, as members of the elites are endowed with initial resources that we postulate are enough for survival.

group) are in proportion λ of the population, with political power β, and workers are in proportion $1 - \lambda$, with political power $1 - \beta$. Let the bourgeois members of the elites (with cultural trait $i = 1$) be in fraction q of the total size of the elites λ, and let the aristocracy be in fraction $1 - q$.

All agents have a technology mapping labor effort one-to-one into private consumption goods. Let a and a^i denote, respectively, the effort of workers and elites with trait $i = 1, 2$. Let p, the policy choice, represent the tax rate on workers' output, a. Let T denote the lump-sum fiscal transfer received by each member of the elites, by budget balance. Let $\bar{c}$ denote the subsistence level required for survival. Recall that workers do not have initial endowments, while all members of the elites—both the aristocrats and the bourgeois—have an endowment $s > \bar{c}$. The per capita fiscal transfer to the members of the elites is set to balance the budget of the fiscal institutions: $T =: ((1 - \lambda)/\lambda)pa$. Preferences are represented by the following utility functions, respectively, for workers and elites:

$$u(a, p) = u(a(1 - p)) + \theta v(1 - a),$$
$$u^i(a^i, T, p) = u(a^i + s + T) + \theta^i v(1 - a^i), \ i = 1, 2.$$

We assume that the aristocrats have extreme preferences for leisure, $\theta^2 > u'(s)/v'(1) > 1 = \theta$, so that they never work, $a^2 = 0$. The bourgeois have instead the same preferences as workers, $\theta^1 = \theta$. A crucial aspect of this society consists in the fact that the labor effort exerted by workers, $a(p)$, is nonmonotonic in the tax rate p, depending on whether the survival constraint is binding. When the tax p is smaller than a cutoff $\hat{p}$, the survival constraint is not binding, and $a(p)$ is decreasing in p because of the disincentive effects of the tax rate on effort. But when instead $p > \hat{p}$, the survival constraint is binding, $a(p) = \bar{c}/(1 - p)$, and workers' labor effort increases with p to maintain survival. We call this the autocratic extractive regime.

The societal equilibrium policy $p(\beta, q)$, and the societal commitment policy $p^{\text{com}}(\beta, q)$ are illustrated in figure 3. When the institutional weight of the elites is high enough, the fiscal authorities, with or without commitment, would choose the extractive regime and tax the workers to a level that forces them to survival. Whether in this regime the commitment problem induces a tax that is too low or too high—that is, whether $P(\beta, q) = p^{\text{com}}(\beta, q) - p(\beta, q)$ is >0 or <0—depends on the balance of two effects. On one hand, as we noted, at survival, higher taxes increase the effort of workers. A fiscal authority lacking commitment would not internalize this effect, inducing a societal equilibrium policy $p(\beta, q)$ lower than the societal equilibrium policy with commitment policy $p^{\text{com}}(\beta, q)$. On the other hand, in this regime, it is also the case that taxes p cause a distortion on workers' welfare that tends to make $p(\beta, q)$ too high compared with $p^{\text{com}}(\beta, q)$ and is not internalized at the societal equilibrium.

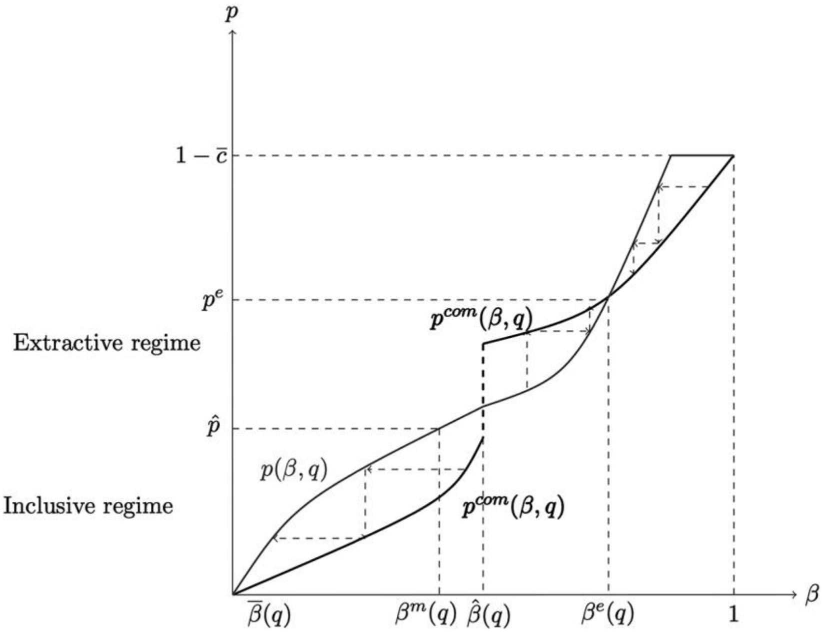


FIG. 3.—Societal equilibrium policies and institutional dynamics.

When β is very high, this distortionary effect on the workers' welfare dominates. But as β decreases, the first effect tends to dominate and $p^{\text{com}}(\beta, q)$ and $p(\beta, q)$ cross, as in figure 3. When β is sufficiently low, society is then out of the extractive regime and into an *inclusive regime*, where workers are not at survival and a fiscal authority with commitment would internalize the negative effect of taxes on the tax base and $p^{\text{com}}(\beta, q) < p(\beta, q)$.

The institutional dynamics, fixing a cultural distribution $0 < q < 1$, depends on the initial condition. For all initial values $\beta_0 > \hat{\beta}(q)$, the dynamics converge to a unique steady state $\beta = \beta^e(q)$, and the society ends up in an autocratic extractive state with low political representation of the workers who are maintained at their survival constraint by extractive taxation on the part of the elites.[55] Conversely, for initial values $\bar{\beta} \leq \beta_0 < \hat{\beta}(q)$, the institutional dynamics are very different. The weight of the elites on

[55] Interestingly in this steady state, higher taxation may actually increase the efficiency of the rent extraction process, as the survival constraint prevents the traditional disincentives on labor supply to kick in. This local effect is arguably instrumental in maintaining such an extractive regime for workers. This is reminiscent of an argument in Clark (2007, chap. 2), suggesting that policies that would otherwise appear as having inefficiency costs in a nonextractive world on the contrary may find some efficiency rationale under extractive conditions.

the institutional setting converge to $\beta = \bar{\beta}$,[56] with no taxation, in an inclusive regime.[57] Interestingly, for initial conditions β_0 between $\beta^m(q)$ and $\hat{\beta}(q)$, the society will move away from the extractive into the inclusive steady state.

The dynamics of cultural evolution within the elite are driven by the relative incentives to socialization $\Delta V^2(p)/\Delta V^1(p)$, which are generally decreasing in p. Indeed, aristocratic norms are more likely to be transmitted than those of the bourgeoisie the larger the rents of the elites. Since equilibrium taxation is a decreasing function of the institutional weight of workers, $1 - \beta$, the more fiscal authority the workers possess in society, the larger the diffusion of (the norms of) the bourgeoisie inside the elite and hence in society.

The joint dynamics of culture and institutions in this society will in general be nonergodic: which stationary state they will converge to in the long run depends on initial conditions (see phase diagram in fig. 4). There are two types of stationary states of the dynamical system: the autocratic extractive state $E = (\beta^e, q^e)$ and (a continuum of equivalent) democratic inclusive steady states, where workers are not taxed, in the interval AB, with $\beta \leq \beta^*$ and $q = q^*$.[58]

Democratization, intended as a transition away from the extractive institutions, is not a necessary feature of the dynamics in this society, as higher taxes do not decrease the fiscal rents of the elites when workers are at or around the survival constraint. Autocratic extractive institutions are not necessarily undermined by their own inefficiency and could be supported in the long run. Whether they are or whether the dynamics transition away depends on the political control the elites exert on the fiscal authority in society but also on their distribution by cultural trait (i.e., the relative size of the bourgeoisie, which is partly aligned with workers' interests). The basin of attraction of the extractive state (β^e, q^e) is the whole region above the red $\hat{\beta}(q)$ line in figure 4.[59]

When the initial institutional setup ensures enough control on fiscal authority on the part of the workers, the dynamics will tend to transition away from the extractive stationary state. But a larger bourgeoisie at the initial conditions also favors the transition away from this state. Formally, the basin of attraction of the democratic inclusive stationary states comprises the whole region (strictly) below the $\hat{\beta}(q)$ line. It is smaller in β for higher q, and it is also larger in q for smaller β. In this region, the dynamics of culture and institutions display a transition away from an autocratic

[56] It stays constant at any value $\beta \leq \bar{\beta}$.

[57] The dynamics from $\beta_0 = \hat{\beta}(q)$ are undetermined.

[58] Cultural complementarities—whereby, e.g., socialization is costlier for minorities—would naturally produce in this example limit distributions, where the elites are composed of only bourgeois (with inclusive institutions) and only aristocrats (with extractive institutions).

[59] It is of interest to note that culture and institutions are substitutes at the extractive state (β^e, q^e), and hence oscillatory dynamics cannot be ruled out from any initial conditions in this region.

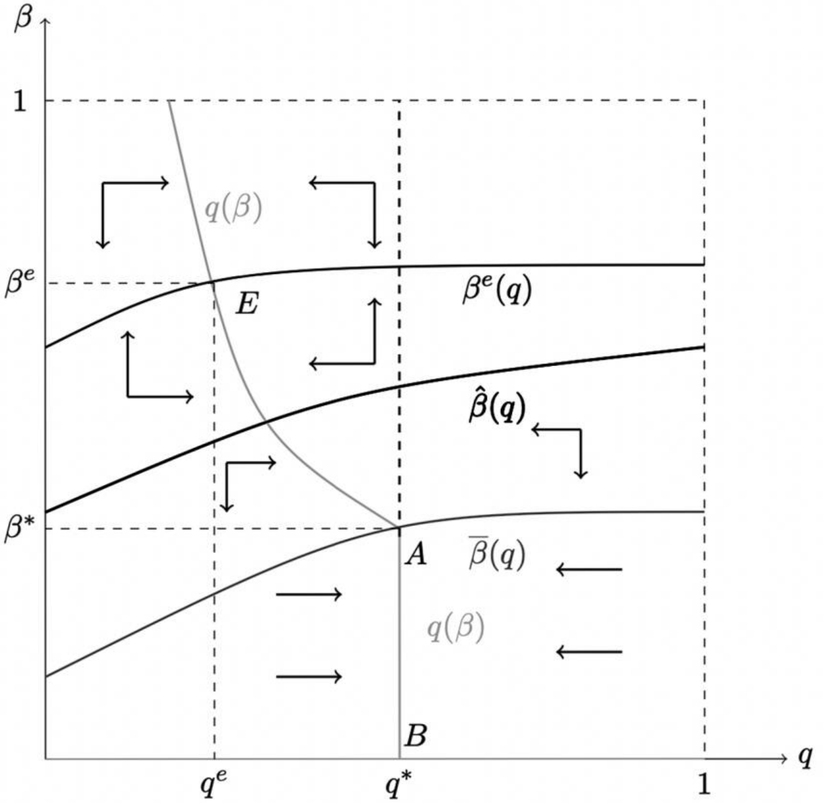


FIG. 4.—Transition away from autocratic extractive states: phase diagram.

extractive society, where workers are taxed and kept at survival, into an inclusive society, where workers consume over and above survival. Furthermore, once an inclusive society is reached, the elites delegate political power to the workers, and taxes decrease. Along this dynamic path, the size of the bourgeoisie grows monotonically.[60]

In this society, the democratic inclusive stationary states are constant in q, and the cultural multiplier is 0 (i.e., there are no effects on β both from any initial condition and at the stationary state; see proposition 8). However, for $q_0 < q^*$, given that the bourgeoisie works and hence—contrary

[60] Several further interesting qualitative implications of the analysis are displayed in the phase diagram. For instance, the relative importance of culture versus institutions as a factor to escape the basin of attraction of the extractive equilibrium can be mapped into the structure of the model (and ultimately into the parameter space). Indeed, when $\hat{\beta}(q)$ is relatively flat, a negative shock on β—e.g., a democratizing shock—is more likely to trigger a path toward an inclusive equilibrium than a positive shock on q—e.g., a bourgeois cultural revolution—and vice versa when the $\hat{\beta}$ has a steeper slope in q.

to the aristocratic leisure class—is productive, the multiplier of aggregate income is positive.

Nonlinearities and causal identification.—The joint evolution of culture and institutions—in general and in particular in the society of the previous section—has some highly nonlinear components. These nonlinearities open up a relevant role for structural models in empirical studies of the determinants of historical phenomena beyond the standard causal identification analysis. We illustrate this with the following conceptual experiment, relying on the phase diagram in figure 4.

Consider three countries with different initial conditions in terms of culture, denoted 1, 2, and 3, as reported in the phase diagram in figure 5. Consider an exogenous shock to institutional setup of the socioeconomic environment affecting the dynamical system, shifting the $\hat{\beta}(q)$ line up to the dashed $\widehat{\beta'}(q)$ line. Since it enlarges the basin of attraction of the inclusive democratic stationary states, this shock can be interpreted as an

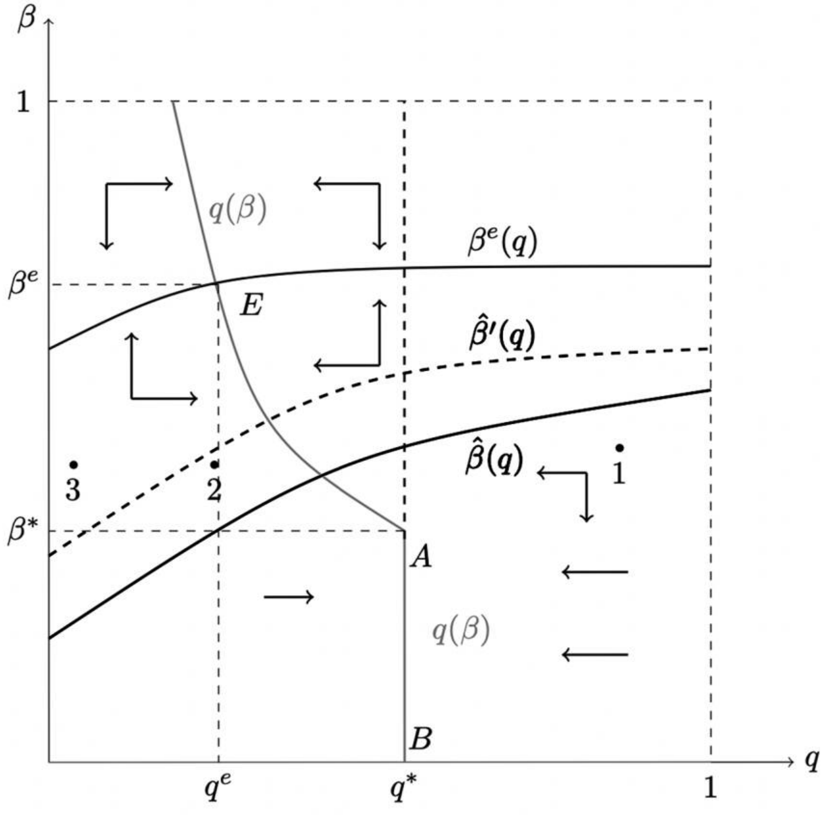


Fig. 5.—(Local) average treatment effects of institutional change.

exogenous push toward democratization. The shock has no qualitative effect on the dynamics of country 1 and 3, the first converging to the autocratic extractive stationary state E and the other to an inclusive democratic state in the segment AB. The shock, on the other hand, leads country 2 toward democratization, converging to a state in AB rather than to E. In the terminology of causal analysis, countries like 1 and 3 are called, respectively, *always takers* and *never takers*; countries like 2 are instead *compliers*. An empirical cross-sectional exercise estimating the effects of this shock on a long-run variable linked to culture and institutions would identify the average treatment effect limited to compliers; that is the local average treatment effect, which will generally differ from the average treatment effect when countries are heterogeneous. In this example, the structural model (represented by the phase diagram) suggests the existence of an initial threshold determining the long-run dynamic path of societies that, when disregarded, induces the econometrician to a misleading interpretation of the effects of an institutional shock to the socioeconomic environment.[61]

B. Civic Capital and Institutions

In this section, we study the dynamics of civic culture in civil society. We model civic culture as a cultural trait that induces actions that may have beneficial societal effects in that they favor the spread of civic capital.[62] In particular, we study how the spread of civic capital in civil society depends on the relative political power of civil society with respect to elites. In this context, we ask under which conditions a loss of political power of the elites, interpreted as a process of democratization, may endogenously induce a decline of civic capital in society, undermining its impact on economic activity. This phenomenon may illustrate a mechanism behind the relative failure of various recent political experiments in democratization, from post-Soviet Russia to Iraq and Afghanistan, highlighting the role of culture—and civic capital in particular—in mediating democratization processes. Along these lines, for example, Guiso and Pinotti (2012) document a decline in civic capital in the south of Italy after the political enfranchisement following the electoral reform of 1912, and Berman (1997) shows how a strong powerful civil society in the Weimar republic lacked the civic capital necessary to its survival.[63]

[61] See Bisin and Moro (2021) for a discussion of these aspects in the context of various persistence studies in the literature and Casey and Klemp (2021) for a bias correction method to instrumental variable estimators in related contexts.

[62] See also Ticchi, Verdier, and Vindigni (2013) and Besley and Persson (2019) for specific analyses of the interactions between political culture and political institutions.

[63] Relatedly, there is also evidence of a democratic backsliding in West Africa, possibly related to a weakening sense of civic duty, as measured by the Afrobarometer (Gyimah-Boadi 2021); see also Fatton (1995).

In this society, both members of the elites and workers are endowed with the same technology that transforms labor effort into private consumption goods. Fiscal institutions collect lump-sum taxes to finance the provision of a public good, whose consumption is valued by both groups. The provision of the public good, however, creates opportunities for corruption that exclusively benefit the elites. We can therefore think of the elites as a caste of bureaucrats. Workers are culturally heterogeneous, in that only a fraction of them have preferences shaped by civic culture. Civic culture motivates workers to exert a participation effort, complementary to the provision of the public good, as well as a monitoring effort to fight corruption. *Civic participation* involves, for example, contributing privately to public goods, creating social associations, and volunteering in social activities. *Civic control* creates transparency by monitoring the government in its public good provision process. Both civic participation and control have consequences that are costly to the elites. Public good provision is therefore associated with different externalities on society. On the one hand, it stimulates the civic participation of a fraction of the workers, a positive externality on society as a whole. On the other hand, public good provision induces corruption and the reaction of a fraction of workers against it, a positive externality on workers with no civic capital and a negative externality on the elite.

Institutions lack commitment; that is, fiscal authorities choose lump-sum taxes to finance public good provision without internalizing the effects of civic capital in society. Public good provision can be larger or smaller at equilibrium than the efficient level, depending on whether the positive or the negative externalities of civic capital dominate. As a consequence, the institutional dynamics lead to a stationary balanced allocation of power between workers and the elite.

Most interestingly, in this society, culture and institutions may act as substitutes. Indeed, on the one hand, the incentives to transmit civic culture are generally increasing in the political representation of workers in society; on the other hand, the larger the spread is of civic capital in the population of workers, the smaller the incentives are to design institutional changes devolving power to the workers, as the beneficial effects of civic capital are already present. In this society, therefore, an exogenous institutional change that endows with more political power citizens could see its effects mitigated by the induced cultural dynamics associated with the spread of civic capital in the population.

Accordingly with the notation in section II, let workers (the culturally heterogeneous group) be a fraction λ of the population with political power β, and let elites be a fraction $1 - \lambda$ with political power $1 - \beta$. Also, workers with a civic culture are (with cultural trait $i = 1$) in fraction q of the total size of workers λ; workers without a civic culture are in fraction $1 - q$. All workers and members of the elites are endowed with a fixed amount of resources, $s > 0$. Lump-sum taxes are raised to finance public expenditures, g. In the process of providing for a public good, a fraction $\mu > 0$ of public expenditures leaks into corruption, generating diverted rents $T = \mu g$ that exclusively benefit the members of the elites. The residual share of public good expenditures is used to provide the public good, $G = (1 - \mu)g$. Workers can exert two types of efforts, civic participation and civic control. Let the participation (control) effort of workers of type i be denoted e^i (a^i). Societal civic participation effort is then $E = \lambda[q \cdot e^1 + (1 - q) \cdot e^2]$, while societal civic control effort is $A = \lambda[q \cdot a^1 + (1 - q) \cdot a^2]$. Societal civic participation effort E produces a society-wide externality that augments each individual's endowment by $\kappa \cdot E$, $\kappa > 0$. Societal civic control effort A increases the transaction costs associated with corruption activities: the consumption associated with T units of diverted rents is $(1 - \theta A)T$, with $0 < \theta < 1$.[64] The government policy is total public expenditures g, financed by lump-sum taxes in the same amount.

The preferences of workers are as follows:

$$u^1(c^1, G, a^1, e^1, T) = c^1 + v(G) - (\alpha \cdot T)(1 - a^1) - C(a^1) + G \cdot e^1 - \Phi(e^1),$$
$$u^2(c^2, G, a^2, e^2, T) = c^2 + v(G) - C(a^2) - \Phi(e^2),$$

where $c^i + v(G)$ is the direct utility of private consumption and the public good for a worker with trait i, $C(a^I)$ is the utility cost of undertaking civic control for a worker with trait i, $\Phi(e^I)$ is the disutility cost of civic participation for a worker with trait i, $-(\alpha \cdot T)(1 - a^1))$ is the intrinsic motivation for civic control of a civic worker, $G \cdot e^1$ is the intrinsic motivation to contribute e^1 to civic participation of a civic worker,[65] while members of the elites have preferences over consumption and the public good:

$$u(c, G) = c + v(G).$$

Policy choice $p = g$ depends on the workers' efforts only through $\theta \cdot A$ and $\kappa \cdot E$. Therefore, since the contribution of each worker effort to societal efforts E and A is negligible, workers with no civic culture always choose not to exert any effort, $a^2 = e^2 = 0$, and workers with civic culture contribute according to their intrinsic motivations. In fact, since both G and T increase in g, e^1 and a^1 also increase in g.

Under some reasonable regularity conditions, $p(\beta, q)$ and $p^{com}(\beta, q)$ are as in figure 8: downward sloping in the political power of workers, β, for any q. More specifically, when workers with civic culture are less in favor of large public expenditures than the elite, an increase in β would tend to reduce the size of the public expenditures at both the societal equilibrium, $p(\beta, q)$, and the societal commitment, $p^{com}(\beta, q)$. For the same reason, at a given value of β, an increase in the fraction of workers with civic culture, q,

[64] Effort costs are normalized so that $\theta A < 1$.
[65] See the online appendix for details, assumptions, and functional forms.

would have the same effect on public expenditures. Most importantly, $p(\beta, q)$ crosses $p^{com}(\beta, q)$ from above at some interior point $\hat{\beta}(q)$. Indeed, $p^{com}(\beta, q)$ is the policy choice once all externalities in society are internalized. But the negative externality, via $\theta \cdot A$, is born out only by elite members, while the positive externality, via E, is enjoyed by the whole society. As a consequence, when the political power of the elite is large (i.e., β small), internalizing the negative externality dominates the society's political objectives and $p^{com}(\beta, q) < p(\beta, q)$. Conversely, when the weight of the elite is small, internalizing the positive externality dominates, and consequently $p^{com}(\beta, q) > p(\beta, q)$. For all initial values β_0, the institutional dynamics converge to a unique steady state $\beta = \hat{\beta}(q)$ and political power is shared between the workers and the elite (see fig. 6).

Importantly, in this society, then the political power of workers at the stationary state, $\hat{\beta}(q)$, is decreasing in the predominance of civic capital, q. This is because societal civic control A increases in q and A substitutes for formal political power. More in detail, at the stationary state $\hat{\beta}(q)$,

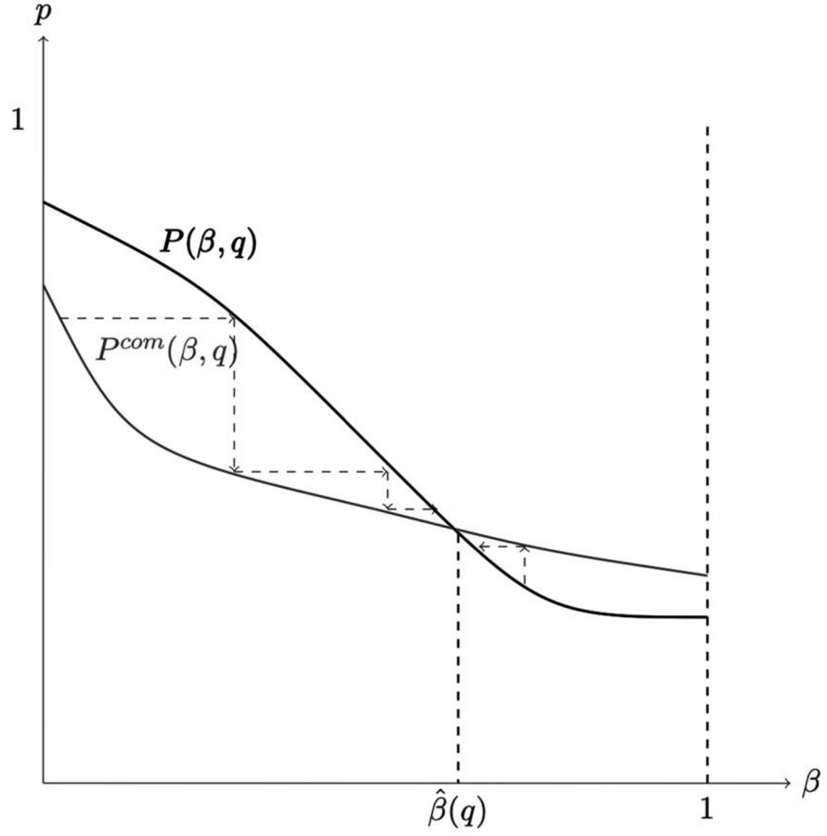


FIG. 6.—Civic capital and institutions: institutional dynamics.

$p^{\mathrm{com}}(\hat{\beta}, q) = p(\hat{\beta}, q)$, and the positive and negative externalities associated with public expenditures balance out at the margin in the government objective function. An increase in q would lead to fewer public expenditures, as workers with civic culture are more concerned than the rest of society with corruption. To restore the equilibrium, institutional dynamics move then in the direction of reintroducing larger public expenditures and hence of reducing the political power of workers, β.

The cultural dynamics within workers are determined by the relative incentives to transmit civic culture, $\Delta V^1(p)/\Delta V^2(p)$, as they depend on the equilibrium policy instrument p. When civic participation e^1 is less sensitive to public good provision than civic monitoring a^1, $\Delta V^1(p)/\Delta V^2(p)$ is decreasing in p. As the societal equilibrium $p(\beta, q)$ is itself a decreasing function of β and q, the relative incentives to transmit civic culture increase with both β and q in society. As a consequence, $q(\beta)$ is upward sloping in β: the formal delegation of power to the workers tends to induce a larger diffusion of civic capital between workers (see fig. 7).

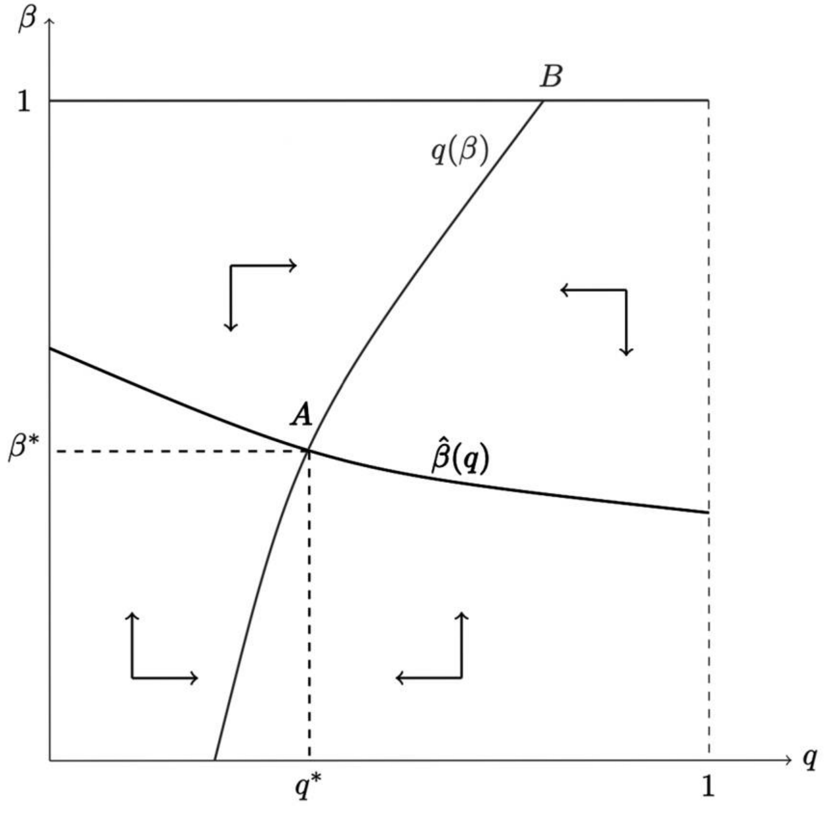


FIG. 7.—Civic capital and institutions: phase diagram.

The joint evolution of culture and institutions is also illustrated in figure 7. The stationary state of the joint dynamics is (β^*, q^*). At (β^*, q^*), the two manifolds $\beta(q)$ and $q(\beta)$ have slopes of opposite signs. As a consequence, culture and institutions are substitutes in this society, and the cultural multiplier is negative (see proposition 8): the effect of an exogenous shock that changes the political power of workers in some direction would be mitigated by the ensuing cultural dynamics. Figure 8 describes the effects of an increase in the coefficient κ, which, other things equal, increases the positive externality associated with civic participation E. A change in κ triggers a higher demand for public expenditures and therefore some institutional dynamics biased against the workers' group. This institutional change in turn reduces the relative incentives to transmit civic culture and leads to a reduction of q. As civic capital is reduced, there is less civic control effort against corruption in society. This in turn calls for some institutional change returning some formal power to workers, therefore mitigating the initial institutional impact of the shock to κ.

Interestingly, depending on the relative speeds of the dynamics of culture and institutions, the dynamics of adjustment to the shock may not be monotonic. Suppose, for instance, that institutions adjust much faster than culture, so that the adjustment dynamics lies on $\hat{\beta}(q)$. In this case, the shock on κ, after having induced β to jump downward (with q constant at q^*), has q decrease and β increase along the adjustment path.

V. Conclusions and Implications for Empirical Studies

In this paper, we develop theoretical and empirical tools for the analysis of the effects of culture and institutions on economic variables of interest, notably, long-run economic activity. In our view, the theoretical model we develop, the concepts of cultural and institutional multipliers we introduce, and the examples we construct to study the political economy of elites and civic society provide some needed structure complementing the role of causal analysis in persistence studies with regard to interesting historical phenomena.

Depending on whether culture and institutions are dynamic complements or substitutes, exogenous historical shocks propagating over the joint dynamics induced by institutions and culture may have magnified or mitigated effects on long-run socioeconomic outcomes. This type of analysis identifies the extent of the comparative dynamics bias that is generated by neglecting one of the two dynamics, when the other one is affected by an exogenous shock. We surmise that this is of first-order importance when studying the long-run effects of historical shocks. Along these lines, while no empirical study has yet attempted to estimate the size of

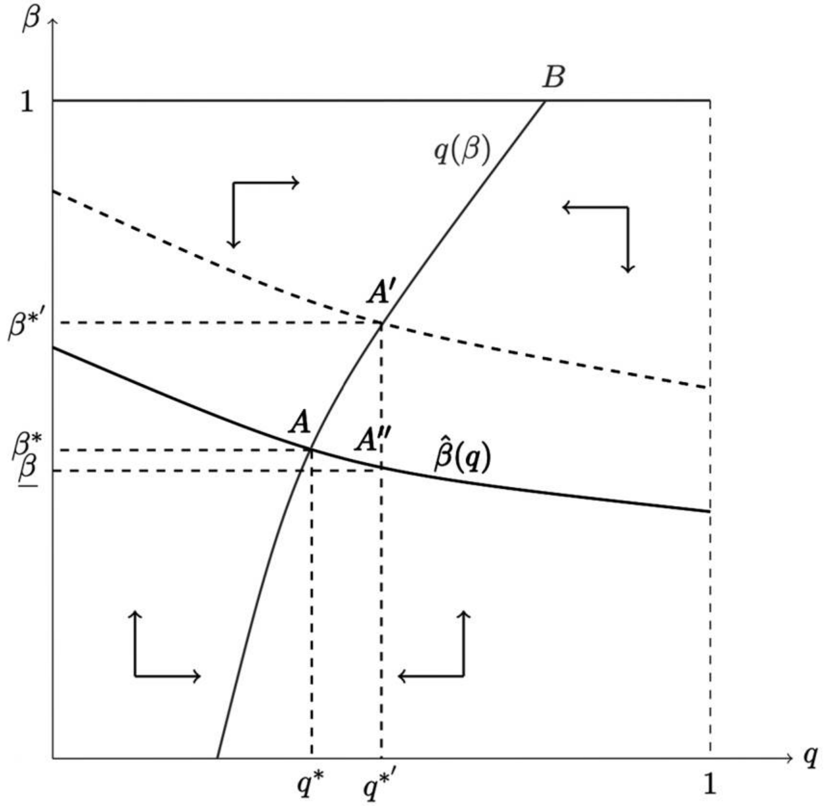


FIG. 8.—Civic capital and institutions: comparative dynamics.

the cultural and the institutional multipliers in the context of persistence studies, several papers provide explicit quantitative evidence about their sign; that is, they document whether culture and institutions acted as complements or substitutes. Lowes et al. (2017) find evidence for substitution along the development of the Kuba Kingdom in Central Africa in the seventeenth century. Lowes and Monteiro (2017) also find substitution in their study of the rubber extraction system in the Congo Free State during the colonial era. On the other hand, Dell's (2010) study of forced mining labor in Peru and Bolivia in the sixteenth century provides suggestive evidence of complementarity.

Our approach also highlights that in general the joint evolution of culture and institutions has some highly nonlinear components. Nonlinearities are at the root of the interesting dynamics we illustrate in the specific example societies in section IV, from sensitivity of equilibrium trajectories to initial conditions to thresholds effects and nonmonotonicity of cultural and institutional changes over transition paths. These phenomena indeed appear quite consistent with the diversity of development trajectories encountered across the world and in time. Most importantly, they highlight the role that culture might have in mediating the relationship between institutions and economic activity.

Appendix A

Supplementary Analysis

We develop here a detailed analysis of several selected topics introduced in the text.

A1. The Social Welfare Function

We here propose microfoundations for the social welfare function $W(\beta, p)$ introduced in section II—equation (1)—to characterize the political mechanism through which collective policy choices are undertaken in the society.

The first microfoundation closely follows Persson and Tabellini (2000) and describes a simple probabilistic voting model in which two rent-seeking candidates maximize their probability of election and compete for electoral support. The second microfoundation is a simple extension with two parties maximizing their plurality in a context of multidistricts elections with a biased institutional structure of voting representation.

A1.1. The Social Welfare Function

Consider two candidates A and B running for an election. Elected officials get a large fixed salary or attain an ego rent R, which is exogenous and fixed. Each candidate commits to an electoral platform to maximize his probability of winning the election and his chance to earn R.

The timing of the game is as follows: in stage 1, candidates announce their policy platforms p_{A}, p_{B} and commit to these if elected; in stage 2, citizens vote; and in stage 3, votes are counted, and the candidate with more votes is elected.

The total voter population is normalized to $L = 1$. Voters are distributed across two political groups: a culturally homogeneous group h (of size λ) and a culturally heterogeneous group m composed of two types of voters: those with cultural trait 1 with size $(1 - \lambda)q$ and those with cultural trait 2 with size $(1 - \lambda)(1 - q)$. The two political groups are differentiated by their electoral franchise: a fraction τ_{h} (τ_{m}) of voters of the homogeneous group (heterogeneous group) is allowed to vote.

Voters care about both policy and candidate characteristics, and they have dispersed (subjective) preferences over candidate characteristics. More precisely, the relative preference of voter i in group $k \in \{\mathrm{h}, \mathrm{m}\}$ for candidate A is represented by the realization of an idiosyncratic random variable ϵ_i^k and a candidate A–specific random variable δ.

Specifically, a voter i in group h prefers candidate A to candidate B if and only if $U(p^{\mathrm{A}}) + \epsilon_i^{\mathrm{h}} + \delta \geq U(p^{\mathrm{B}})$, while a voter i in group m of cultural type $j \in \{1, 2\}$ prefers candidate A if and only if $U_j(p^{\mathrm{A}}) + \epsilon_i^{\mathrm{m}} + \delta \geq U_j(p^{\mathrm{B}})$. $U(p)$ and $U_j(p) j \in \{1, 2\}$ are the indirect utility functions of the relevant voter with respect to the platform policy p. These functions are assumed to be continuous and strictly concave in p.

Following Persson and Tabellini (2000), we assume that the idiosyncratic component ϵ_i^k is uniformly distributed for individuals of group $k \in \{\mathrm{h}, \mathrm{m}\}$ on $[b^k - (1/2\phi^k), b^k + (1/2\phi^k)]$, while candidate A–specific random component δ is uniformly distributed on $[-(1/2\psi), (1/2\psi)]$.

The indifferent voter i in group h and in group m (for the two types $j = 1, 2$) is characterized by an idiosyncratic cutoff for the realization of ϵ_i^k. This is given by, respectively,

$$\begin{aligned}
\epsilon_i^{\mathrm{h}} &= U(p^{\mathrm{B}}) - U(p^{\mathrm{A}}) - \delta \text{ and} \\
\epsilon_i^{\mathrm{m}} &= U_j(p^{\mathrm{B}}) - U_j(p^{\mathrm{A}}) - \delta \text{ for } j \in \{1, 2\}.
\end{aligned}$$

The fraction of votes for candidate A in each category of voters is then easily obtained as

$$\begin{aligned}
\pi^{\mathrm{hA}} &= \frac{1}{2} + \phi^{\mathrm{h}} b^{\mathrm{h}} + \phi^{\mathrm{h}}[U(p^{\mathrm{A}}) - U(p^{\mathrm{B}}) + \delta], \\
\pi_j^{\mathrm{mA}} &= \frac{1}{2} + \phi^{\mathrm{m}} b^{\mathrm{m}} + \phi^{\mathrm{m}}\left[U_j(p^{\mathrm{A}}) - U_j(p^{\mathrm{B}}) + \delta\right], j \in \{1, 2\}.
\end{aligned}$$

The total amount of votes for candidate A is

$$n^{\mathrm{A}} = \lambda\tau_{\mathrm{h}}\pi^{\mathrm{hA}} + (1 - \lambda)\tau_{\mathrm{m}}[q\pi_1^{\mathrm{mA}} + (1 - q)\pi_2^{\mathrm{mA}}].$$

As well, the probability for candidate A to win the elections is

$$P^{\mathrm{A}} = \Pr\left(n^{\mathrm{A}} > \frac{\lambda\tau_{\mathrm{h}} + (1 - \lambda)\tau_{\mathrm{m}}}{2}\right),$$

or after substitution of n^{A} and manipulation,

$$\begin{aligned}
P^{\mathrm{A}} = \Pr\Bigg(\ \delta > &-\frac{\lambda\tau_{\mathrm{h}}\phi^{\mathrm{h}} b^{\mathrm{h}} + (1 - \lambda)\tau_{\mathrm{m}}\phi^{\mathrm{m}} b^{\mathrm{m}}}{\lambda\tau_{\mathrm{h}}\phi^{\mathrm{h}} + (1 - \lambda)\tau_{\mathrm{m}}\phi^{\mathrm{m}}} \\
&- \frac{\lambda\tau_{\mathrm{h}}\phi^{\mathrm{h}}}{\lambda\tau_{\mathrm{h}}\phi^{\mathrm{h}} + (1 - \lambda)\tau_{\mathrm{m}}\phi^{\mathrm{m}}}[U(p^{\mathrm{A}}) - U(p^{\mathrm{B}})] \\
&- \frac{(1 - \lambda)\tau_{\mathrm{m}}\phi^{\mathrm{m}}}{\lambda\tau_{\mathrm{h}}\phi^{\mathrm{h}} + (1 - \lambda)\tau_{\mathrm{m}}\phi^{\mathrm{m}}}[q[U_1(p^{\mathrm{A}}) - U_1(p^{\mathrm{B}})] + (1 - q)[U_2(p^{\mathrm{A}}) - U_2(p^{\mathrm{B}})]]\Bigg),
\end{aligned}$$

or integrating on δ,

$$\begin{aligned}
P^{\mathrm{A}} = P(p^{\mathrm{A}}, p^{\mathrm{B}}) &= \frac{1}{2} + \psi\big((1 - \beta)b^{\mathrm{h}} + \beta b^{\mathrm{m}}\big) + \psi[(1 - \beta)[U(p^{\mathrm{A}}) - U(p^{\mathrm{B}})] \\
&\quad + \beta[q[U_1(p^{\mathrm{A}}) - U_1(p^{\mathrm{B}})] + (1 - q)[U_2(p^{\mathrm{A}}) - U_2(p^{\mathrm{B}})]]],
\end{aligned}$$

with

$$\beta = \frac{(1-\lambda)\tau_{\mathrm{m}}\phi^{\mathrm{m}}}{\lambda\tau_{\mathrm{h}}\phi^{\mathrm{h}} + (1-\lambda)\tau_{\mathrm{m}}\phi^{\mathrm{m}}}. \tag{1}$$

Party A maximizes his expected political rent $P(p^{\mathrm{A}}, p^{\mathrm{B}}) \cdot R$ on p^{A}, while party B maximizes his expected rent $(1 - P(p^{\mathrm{A}}, p^{\mathrm{B}})) \cdot R$ on p^{B}. It is simple to see that the Nash equilibrium policy platforms of the two candidates converge to $p^{\mathrm{A}} = p^{\mathrm{B}} = p^*$ such that

$$p^* = \arg\max_p W(p, \beta) = (1-\beta) \cdot U(p) + \beta \cdot [qU_1(p) + (1-q)U_2(p)].$$

Hence, our formulation of the policy choice in the main text as reflecting the maximization of an objective function $W(p, \beta)$ can be rationalized as the outcome of the political competition between two candidates maximizing their expected rents in a probabilistic voting context. From (1), the institutional parameter β characterizing the relative political institutional power of the heterogeneous political group m directly connects to the relative institutional voting franchise $\tau_{\mathrm{m}}/\tau_{\mathrm{h}}$.

A1.2. Party Competition in District Voting

We consider the previous model of probabilistic voting with two parties A and B competing for election. We allow here voters to be distributed across a continuum of districts $z \in [0, 1]$ with different degrees of political representation. We also assume that parties maximize their plurality (i.e., share of seats in the parliament).

For this, let $\lambda(z)$ be the fraction of individuals of the homogeneous political group h in district z. Without loss of generality, assume that $\lambda'(z) > 0$. We have $\int_0^1 \lambda(z)dz = \lambda$. Also assume that $\lambda(0) = 0$ and $\lambda(1) = 1$. We define $\bar{z} < 1$ the district such that $\lambda(\bar{z}) = \lambda$. In the heterogeneous group, we consider for simplicity that the share of voters of type 1 and type 2 is uniformly distributed across districts (i.e., the fraction of individuals of type 1 of the heterogeneous group located in district z is $q(z) = q$ for all $z \in [0, 1]$.

Because of administrative and political dimensions, districts differ in their political voting weights $\gamma(z)$. To fix ideas, consider that all districts $z < \bar{z}$ are institutionally underrepresented and members in such districts have a voting weight $\gamma(z) = 1 - v$. Conversely, all districts with $z > \bar{z}$ are institutionally overrepresented and members in such districts have a voting weight $\gamma(z) = (1 + v)$, with $v \in [-1, 1]$, reflecting the degree of biased representation across the two types of districts.

Finally, consider that parties maximize plurality. Then, total votes for party A is written as

$$n^{\mathrm{A}} = \left[\int_0^1 \lambda(z)\gamma(z)\,dz\right]\pi^{\mathrm{hA}} + \left[\int_0^1 (1-\lambda(z))\gamma(z)\,dz\right][q\pi_1^{\mathrm{mA}} + (1-q)\pi_2^{\mathrm{mA}}],$$

or for the specific political weights that we assumed,

$$\begin{aligned} n^{\mathrm{A}} = &\left[(1-v)\int_0^{\bar{z}} \lambda(z)dz + (1+v)\int_{\bar{z}}^1 \lambda(z)\,dz\right]\pi^{\mathrm{hA}} \\ &+ \left[(1-v)\int_0^{\bar{z}} (1-\lambda(z))\,dz + (1+v)\int_{\bar{z}}^1 (1-\lambda(z))\,dz\right][q\pi_1^{\mathrm{mA}} + (1-q)\pi_2^{\mathrm{mA}}]. \end{aligned}$$

Substitution of π^{hA}, π_1^{mA}, and π_2^{mA} and algebraic manipulations provides

$$n_A = [(1-v)\bar{z} + (1+v)(1-\bar{z})] \cdot \left[\begin{array}{c} \Delta(v)\left[\frac{1}{2} + \phi^h b^h + \phi^h[U(p^A) - U(p^B)]\right] \\ +[1-\Delta(v)] \cdot \left[\begin{array}{c} \frac{1}{2} + \phi^m b^m + \phi^m q[U_1(p^A) - U_1(p^B)] \\ +\phi^m(1-q)[U_2(p^A) - U_2(p^B)] \end{array}\right] \end{array}\right],$$

with

$$\Delta(v) = \left[\begin{array}{c} \frac{(1-v)}{(1-v)\bar{z} + (1+v)(1-\bar{z})} \int_0^{\bar{z}} \lambda(z)\,dz \\ + \frac{(1+v)}{(1-v)\bar{z} + (1+v)(1-\bar{z})} \int_{\bar{z}}^1 \lambda(z)\,dz \end{array}\right].$$

Simple computations show that $\Delta'(v)$ is proportional to

$$\begin{aligned} \int_{\bar{z}}^1 \lambda(z)\,dz - \int_0^{\bar{z}} \lambda(z)\,dz - \lambda + 2\bar{z}\lambda &= 2(\bar{z}\lambda - \int_0^{\bar{z}} \lambda(z)\,dz) \\ &= 2(\bar{z}\lambda(\bar{z}) - \int_0^{\bar{z}} \lambda(z)\,dz) > 0, \end{aligned}$$

as $\lambda'(z) > 0$ and $\bar{z} > 0$. Consequently, $\Delta(v)$ is increasing in the district bias v.

Plurality of party A is written as

$$\begin{aligned} \pi_A^T(p^A, p^B) &= \frac{n_A}{(1-v)\bar{z} + (1+v)(1-\bar{z})} \\ &= \Delta(v)\left[\frac{1}{2} + \phi^h b^h + \phi^h[U(p^A) - U(p^B)]\right] \\ &\quad + [1-\Delta(v)]\left[\begin{array}{c} \frac{1}{2} + \phi^m b^m + \phi^m q[U_1(p^A) - U_1(p^B)] \\ +\phi^m(1-q)[U_2(p^A) - U_2(p^B)] \end{array}\right]. \end{aligned}$$

For simplicity, pose $b^h = b^m = 0$; then plurality of party A becomes

$$\begin{aligned} \pi_A^T(p^A, p^B) &= \frac{1}{2} + \Delta(v)\phi^h[U(p^A) - U(p^B)] \\ &\quad + (1-\Delta(v))\phi^m \cdot [q[U_1(p^A) - U_1(p^B)] + (1-q)[U_2(p^A) - U_2(p^B)]]. \end{aligned}$$

Denote

$$\beta(v) = \frac{(1-\Delta(v))\phi^m}{\Delta(v)\phi^h + (1-\Delta(v))\phi^m}.$$

It is easy to see that $\pi_A^T(p^A, p^B)$ is proportional to

$$(1-\beta) \cdot [U(p^A) - U(p^B)] + \beta[q[U_1(p^A) - U_1(p^B)] + (1-q)[U_2(p^A) - U_2(p^B)]],$$

while plurality of party B is simply $\pi_B^T(p^A, p^B) = 1 - \pi_A^T(p^A, p^B)$. It is clear that the Nash equilibrium policy outcome $p^* = p^{A*} = p^{B*}$ emanating from the political competition between parties A and B with biased proportionality across districts is the same as the one chosen by a policy maker maximizing the weighted welfare function of the main text,

$$W(\beta, p) = (1 - \beta)U(p) + \beta \cdot [qU_1(p) + (1 - q)U_2(p)], \tag{2}$$

with a weight structure $\beta = \beta(v)$ being determined by the institutional relative voting bias of districts. Any change in the district bias v (administrative reform, redesign of geographic district contours [i.e., gerrymandering]) implies a change in the political weight β of the objective function in (2).

A2. Slippery Slope Effects

We here develop an analysis of slippery slope effects as a consequence of forward-looking institutional dynamics, introduced in section III.C.1.

Define more specifically the Nash equilibrium behavioral vector of society for a given policy level p as $\mathbf{a}(p, q) = [a(p, q), a^1(p, q), a^2(p, q)]$.[66] The policy maker objective function associated with such equilibrium behavioral response associated with a policy p is rewrittten as

$$\bar{W}(\beta_t, p, q) = W(\beta_t, \mathbf{a}(p, q), p, q)$$

for the current period.[67]

For current institutions β_t, the one-step forward-looking institutional design is obtained from the perfect Nash equilibrium $(\beta^1\ (\beta t),\ \tilde{\beta}^0,\ (\beta t))$ of the following two-period game. In the first stage, initial institutions β_t design institutions β_{t+1},. In the second stage, the institutions β_{t+1} in turn design the next period institutions β'_{t+2} in a myopic way. $(\beta^1\ (\beta t),\ \tilde{\beta}^0,\ (\beta t))$ should then satisfy

$$\beta^1(\beta_t) \in \arg\max_{\beta'} \bar{W}(\beta_t, p(\beta', q), q) + \frac{\delta}{1-\delta} \bar{W}(\beta_t, p(\beta^0(\beta'), q), q),$$

where $\beta^0\ (\beta)$ is the optimal myopic institutions β_{t+2} designed by a given second-stage institution player β, that is,

$$\beta^0(\beta) \in \arg\max_{\beta'} \frac{1}{1-\delta} \bar{W}(\beta, p(\beta', q), q)$$

[66] For simplicity, we assume that conditions hold to ensure the existence of a unique such Nash equilibrium behavioral vector $\mathbf{a}(p, q)$. Formally, it is the fixed point of the following correspondence T, which associates with each behavioral vector $\mathbf{a} = [a, a^1, a^2]$ the behavioral vector

$$T(\mathbf{a}) = \left[\tilde{a}(p, q, \mathbf{a}), \tilde{a}^1(p, q, \mathbf{a}), \tilde{a}^2(p, q, \mathbf{a})\right]$$

such that

$$\tilde{a}(p, q, \mathbf{a}) \in \arg\max_a u(a, p; \mathbf{a}, q),$$

$$\tilde{a}^i(p, q, \mathbf{a}) \in \arg\max_a u^i(a, p; \mathbf{a}, q) \text{ for } i \in \{1, 2\}.$$

[67] We consider the case where $\beta^0(\beta_{t+1}) \in (0, 1)$ is an interior solution. The argument can be appropriately accommodated when $\beta^0(\beta_{t+1}) = 0, 1$.

and

$$\tilde{\beta}^0(\beta_t) = \beta^0\left(\beta^1(\beta_t)\right).$$

One can solve the game by backward induction. In stage 2, the institutional player β_{t+1} is myopic, and given that social welfare evaluated at weights β_{t+1} is highest under policy $p^{\mathrm{com}}(\beta_{t+1}, q)$, the optimal institutional choice β^0_{t+2} is designed to induce the choice $p^{\mathrm{com}}(\beta^1_{t+1}, q)$; that is, $p(\beta^0_{t+2}, q) = p^{\mathrm{com}}(\beta^1_{t+1}, q)$ or $\beta^0_{t+2} = \beta^0(\beta_{t+1})$. Consequently, in stage 1, the optimal choice of institutions by the current institutional framework β_t is given by

$$\beta^1_{t+1} \in \arg\max_{\beta'} \bar{W}(\beta_t, p(\beta', q), q) + \frac{\delta}{1-\delta}\bar{W}(\beta_t, p^{\mathrm{com}}(\beta', q), q).$$

Define the function

$$\Psi(\beta_t, \beta, q) = \bar{W}(\beta_t, p(\beta, q), q) + \frac{\delta}{1-\delta}\bar{W}(\beta_t, p^{\mathrm{com}}(\beta, q), q),$$

and assume that the function $\Psi(\beta_t, \beta, q)$ is smooth and strictly concave in β, so that a first-order approach holds. Formally, we require that the policy functions $p(\beta, q)$ and $p^{\mathrm{com}}\,(\beta, q)$ are smooth and that

$$\bar{W}_{pp}\left(\frac{\partial p}{\partial \beta}\right)^2 + \bar{W}_p\frac{\partial^2 p}{\partial \beta^2} + \frac{\delta}{1-\delta}\bar{W}^{\mathrm{com}}_{pp}\left(\frac{\partial p^{\mathrm{com}}}{\partial \beta}\right)^2 + \bar{W}_p\frac{\partial^2 p^{\mathrm{com}}}{\partial \beta^2} < 0.$$

This will be ensured when the function $\bar{W}(\beta_t, p, q) = W(\beta_t, \mathbf{a}(p, q), p, q)$ is a well-defined smooth enough concave function of the policy level p.[68]

An interior solution $\beta^1_{t+1} = \beta^1(\beta_t) \in (0,1)$ should then satisfy

$$\begin{aligned}\Psi_\beta\left(\beta_t, \beta^1_{t+1}, q\right) &= \bar{W}_p\left(\beta_t, p\left(\beta^1_{t+1}, q\right), q\right)\cdot\left.\frac{\partial p}{\partial \beta}\right|_{\beta^1_{t+1}}\\ &\quad+\frac{\delta}{1-\delta}\bar{W}_p\left(\beta_t, p^{\mathrm{com}}(\beta^1_{t+1}, q), q\right)\cdot\left.\frac{\partial p^{\mathrm{com}}}{\partial \beta}\right|_{\beta^1_{t+1}}\\ &= 0.\end{aligned}$$

To illustrate the impact of the slippery slope argument, let us compare for a given current institutional player β_t the one-step forward-looking institutional

[68] This is ensured when

$$\frac{\partial^2 W}{\partial p^2} < 0$$

and

$$\begin{aligned}\bar{W}_{pp}(\beta_t, p, q) &= \frac{\partial W}{\partial a}\cdot\frac{\partial^2 a}{\partial p^2} + \sum_{i=1,2}\frac{\partial W}{\partial a^i}\cdot\frac{\partial^2 a^i}{\partial p^2}\\ &\quad+\frac{\partial^2 W}{\partial a\partial p}\cdot\frac{\partial a}{\partial p} + 2\sum_{i=1,2}\frac{\partial^2 W}{\partial a^i\partial p}\frac{\partial ai}{\partial p}\\ &\quad+\frac{\partial^2 W}{\partial p^2}\end{aligned}$$

is sufficiently negative.

design $\beta^1(\beta_t)$ with the myopic choice $\beta^0(\beta_t)$. Suppose also that the myopic institutional design path $\beta_{t+1} = \beta^0(\beta_t)$ converges monotonically to the stable institutional steady state $\bar{\beta} \in (0, 1)$, and assume that $p^{\text{com}}(\beta, q)$ is increasing in β. As we know, this steady state is characterized by $p(\bar{\beta}, q) = p^{\text{com}}(\bar{\beta}, q)$, $p(\beta, q)$ is also increasing in β, and $p(\beta, q) \geq p^{\text{com}}(\beta, q)$ if and only if $\beta \leq \bar{\beta}$.

i. Note first that the steady state $\bar{\beta}$ is also a steady state of the one-step forward-looking institutional dynamics design $\beta_{t+1} = \beta^1(\beta_t)$ (i.e., $\bar{\beta} = \beta^1_{t+1}(\bar{\beta})$). Indeed, taking into account the fact that $p(\bar{\beta}, q) = p^{\text{com}}(\bar{\beta}, q)$ and that $\bar{W}_p(\beta, p^{\text{com}}(\beta, q), q) = 0$, we have

$$\begin{aligned}\Psi_\beta(\bar{\beta}, \bar{\beta}, q) &= \bar{W}_p(\bar{\beta}, p(\bar{\beta}, q), q) \cdot \left.\frac{\partial p}{\partial \beta}\right|_{\bar{\beta}} + \frac{\delta}{1-\delta} \bar{W}_p(\bar{\beta}, p^{\text{com}}(\bar{\beta}, q), q) \cdot \left.\frac{\partial p^{\text{com}}}{\partial \beta}\right|_{\bar{\beta}} \\ &= \bar{W}_p(\bar{\beta}, p^{\text{com}}(\bar{\beta}, q), q) \cdot \left[\left.\frac{\partial p}{\partial \beta}\right|_{\bar{\beta}} + \frac{\delta}{1-\delta}\left.\frac{\partial p^{\text{com}}}{\partial \beta}\right|_{\bar{\beta}}\right] = 0.\end{aligned}$$

ii. As well, note that

$$\begin{aligned}\Psi_\beta(\beta_t, \beta_t, q) &= \bar{W}_p(\beta_t, p(\beta_t, q), q) \cdot \left.\frac{\partial p}{\partial \beta}\right|_{\beta_t} + \frac{\delta}{1-\delta} \bar{W}_p(\beta_t, p^{\text{com}}(\beta_t, q), q) \cdot \left.\frac{\partial p^{\text{com}}}{\partial \beta}\right|_{\beta_t} \\ &= \bar{W}_p(\beta_t, p(\beta_t, q), q) \cdot \left.\frac{\partial p}{\partial \beta}\right|_{\beta_t}.\end{aligned}$$

The concavity of $\bar{W}$ in p and the fact that $\bar{W}$ reaches its maximum at $p^{\text{com}}(\beta, q)$, implies that $\Psi_\beta(\beta_t, \beta_t, q) \geq 0$ iff $p(\beta_t, q) \leq p^{\text{com}}(\beta_t, q)$ and $\Psi_\beta(\beta_t, \beta_t, q) \geq 0$ if and only if $\beta_t \leq \bar{\beta}$. The concavity of $\Psi(\beta_t, \beta, q)$ in β therefore implies that $\beta_t \leq \beta^1(\beta_t)$ if and only if $\beta_t \leq \bar{\beta}$.

iii. Finally, note that at point $\beta^0_{t+1} = \beta^0(\beta_t)$, by definition, one has $p^{\text{com}}(\beta^0_{t+1}, q) = p(\beta_t, q)$, and again noting that $\bar{W}_p(\beta_t, p^{\text{com}}(\beta_t, q), q) = 0$, one gets

$$\Psi_\beta(\beta_t, \beta^0(\beta_t), q) = \frac{\delta}{1-\delta} \bar{W}_p(\beta_t, p^{\text{com}}(\beta^0(\beta_t), q), q) \cdot \left.\frac{\partial p^{\text{com}}}{\partial \beta}\right|_{\beta^0(\beta_t)},$$

the sign of which depends on the sign of $\bar{W}_p(\beta_t, p^{\text{com}}(\beta^0(\beta_t), q), q)$. Thus, with the concavity of $\bar{W}$ in p and the fact that $\bar{W}$ reaches its maximum at $p^{\text{com}}(\beta, q)$, one has $\Psi_\beta(\beta_t, \beta^0(\beta_t), q) \geq 0$ if and only if $p^{\text{com}}(\beta^0(\beta_t), q) \leq p^{\text{com}}(\beta_t, q)$ or $\beta^0(\beta_t) \leq \beta_t$. The concavity of $\Psi(\beta_t, \beta, q)$ in β implies that $\beta^0(\beta_t) \leq \beta^1(\beta_t)$ if and only if $\beta^0(\beta_t) \leq \beta_t$.

For all $\beta_t < \bar{\beta}$, we have $\beta_t < \beta^0(\beta_t)$ and consequently $\beta_t < \beta^1(\beta_t) < \beta^0(\beta_t)$, while for $\beta_t > \bar{\beta}$, we have $\beta_t > \beta^0(\beta_t)$ and consequently $\beta^0(\beta_t) < \beta^1(\beta_t) < \beta_t$. From this, we can conclude that the one-step forward-looking institutional change converges to the same steady state $\bar{\beta}$ as the myopic institutional change but at a reduced speed.

A3. Predator-Prey Cycles

We develop here the extension introduced in section III.C.3, with two culturally heterogeneous political groups displaying predator-prey cultural dynamics and limit cycles.

More precisely, consider two culturally heterogeneous political groups denoted E and V of size λ and $1 - \lambda$, respectively. More precisely, in political group E, there are two cultural subgroups denoted E_1 and E_2 in proportions q_E and $1 - q_E$. Similarly, in political group V, there are two cultural subgroups V_1 and V_2 in proportions q_V and $1 - q_V$.

Preferences take the following separable form:

$$U_i^g(a_i, p, A) = u_i^g(a_i, p) + H^g(p, A) \text{ for } g \in \{E, V\}, i \in \{g_1, g_2\} \quad (3)$$

where a_i is the typical social individual action of an individual, p is the policy with $p \in [p_{\min}, p_{\max}]$, and A is an aggregate action index with $A = A(\mathbf{a}, q_E, q_V)$, with $\mathbf{a} = (a_{E_1}, a_{E_2}, a_{V_1}, a_{V_2})$, the vector of individual action types (as in the main text, we concentrate on societal equilibria where all agents of a given cultural type $i \in \{g_1, g_2\}$ in political group $g \in \{E, V\}$ take the same equilibrium action).

The policy objective function is given by

$$\begin{aligned} W = {} & \beta[q_E \cdot U_{E_1}^E(a_{E_1}, p, A) + (1 - q_E) \cdot U_{E_2}^E(a_{E_2}, p, A)] \\ & + (1 - \beta) \cdot [q_V \cdot U_{V_1}^V(a_{V_1}, p, A) + (1 - q_V) \cdot U_{V_2}^V(a_{V_2}, p, A)], \end{aligned}$$

with β the political weight of group E.

Following our general logic, the dynamic system of institutional and cultural change can be written as

$$\dot{\beta} = h[p^{\text{com}}(\beta, q_E, q_V) - p(\beta, q_E, q_V)], \text{ with } h > 0, \quad (4)$$

$$\begin{aligned} \dot{q}_E &= q_E(1 - q_E)S_E(q_E, q_V, p), \\ \dot{q}_V &= q_V(1 - q_V)S_V(q_V, q_E, p). \end{aligned} \quad (5)$$

Equation (4) describes the institutional dynamics with $p(\beta, q_E, q_V)$, the societal equilibrium policy from the policy game in a given period, and $p^{\text{com}}(\beta, q_E, q_V)$, the societal commitment equilibrium policy. Both policies are now depending on the institutional weight β and the cultural composition of the two political groups, as reflected by q_E, q_V. Equations (5) are the cultural replicator equations for the two traits E_1 and V_1, assuming that the cultural transmission process occurs only within political groups. Moreover, because of the separability of preferences in (3), and the fact that the aggregate index A affects in the same way both cultural subgroups within each political group (i.e., $H^g(p, A)$ does not depend on $i \in \{g_1, g_2\}$), one gets a simpler cultural dynamic system:

$$\begin{aligned} \dot{q}_E &= q_E(1 - q_E)S_E(q_E, p), \\ \dot{q}_V &= q_V(1 - q_V)S_V(q_V, p). \end{aligned} \quad (6)$$

In the case of the Bisin Verdier (BV) cultural evolutionary process with quadratic socialization costs, the relative cultural fitness $S_g(q_g, p)$ takes the following form:

$$\begin{aligned} S_g(q_g, p) = {} & (1 - q_g)\left[u_{g_1}^g(a_{g_1}(p), p) - u_{g_1}^g(a_{g_2}(p), p)\right] \\ & - q_g\left[u_{g_2}^g(a_{g_2}(p), p) - u_{g_2}^g(a_{g_1}(p), p)\right]. \end{aligned}$$

For the institutional equation (4), we assume the following:

$$\text{Assumption I}: \ \frac{\partial p}{\partial \beta} > \frac{\partial p^{\text{com}}}{\partial \beta} > 0 \text{ for all } (q_E, q_V) \in [0, 1]^2. \quad (7)$$

Political group 1 is in favor of the policy p, and the institutional dynamic on β is monotonic, converging toward a stable steady state manifold $\beta = \bar{\beta}(q_E, q_V) \in [0, 1]$ for all $(q_E, q_V) \in [0, 1]^2$.

For the cultural transmission equations (6), we assume the following:

$$\text{Assumption C}: \ \frac{\partial S_g(q_g, p)}{\partial q_g} < 0 \text{ for all } p \in P; \ S_g(0, p) > 0 > S_g(1, p), g \in \{E, V\}. \quad (8)$$

This assumption is satisfied under strong enough cultural substitution in the BV model and ensures that for a given value of the policy p, the cultural dynamics converge to an interior steady state $\bar{q}_g(p) \in (0, 1)$. Moreover, we assume the following monotonicity assumption: $\partial S_g(q_g, p)/\partial p > 0$ for all $g \in \{E, V\}$. Essentially, in each political group g = E, V, we label cultural subgroup g_1, the group whose diffusion is favored by a higher value of the policy p. The implicit assumption is that the policy p always promotes a given cultural subgroup inside each political group. These conditions therefore ensure the existence of stationary cultural manifolds $\bar{q}_g(p) \in (0, 1)$ for all p with $\bar{q}'_g(p) > 0$.

We now say that culture and institutions are *cultural substitutes* (*cultural complements*) with respect to trait g_1, with $g \in \{E, V\}$ when $(\partial p^{com}/\partial q_g) - (\partial p/\partial q_g) < 0$ (>0). In other words, an increase in the fraction q_g of trait g_1 promoted by the institutional empowerment of group g leads to a reduction (an increase) in the need of further institutional change in the same direction. Indeed, the commitment issue to be resolved becomes less intense as $(\partial p^{com}/\partial q_g) - (\partial p/\partial q_g) < 0$ (more intense as $(\partial p^{com}/\partial q_g) - (\partial p/\partial q_g) > 0$). This leads to the fact that $\bar{\beta}(q_E, q_V)$ is decreasing (increasing) in q_g.

Now under assumptions I and C, a steady state of the system (4) and (6) is characterized by

$$S_E(q_E, p) = 0\ ,\ S_V(q_V, p) = 0, \beta = \bar{\beta}(q_E, q_V), \text{ and } p(\beta, q_E, q_V) = p,$$

or

$$q_E = \bar{q}_E(p), q_V = \bar{q}_V(p), \text{and } \beta = \bar{\beta}(p) = \bar{\beta}(\bar{q}_E(p), \bar{q}_V(p)) \in [0, 1].$$

Consider then the function

$$\Omega(p) = p - p(\bar{\beta}(p), \bar{q}_E(p), \bar{q}_V(p)).$$

The existence of a steady state of (4) and (6) is ensured when $\Omega(p)$ is continuous, as $\Omega(p_{\min}) \leq 0 \leq \Omega(p_{\max})$), and therefore there exists a value $p^* \in [p_{\min}, p_{\max}]$ such that $\Omega(p^*) = 0$. Such a steady state is characterized by $q_E^* = \bar{q}_E(p^*)$, $q_V^* = \bar{q}_V(p^*)$, and $\beta^* = \bar{\beta}(p^*)$. A sufficient condition for a unique steady state is that $\Omega'(p) < 0$ or

$$\sum_{g \in \{E,V\}} \left(\frac{\partial p}{\partial \beta} \cdot \frac{(\partial p^{com}/\partial q_g) - (\partial p/\partial q_g)}{(\partial p/\partial \beta) - (\partial p^{com}/\partial \beta)} + \frac{\partial p}{\partial q_g} \right) \bar{q}'_g(p) < 1. \quad (9)$$

An interesting special case is the one where institutions are a fast-moving variable compared with culture. In that case, $h \rightarrow \infty$ and β is always characterized by the steady state manifold $\beta = \bar{\beta}(q_E, q_V) \in [0, 1]$ for all $(q_E, q_V) \in [0, 1]^2$. The cultural dynamics are then rewritten as

$$\begin{aligned} \dot{q}_E &= q_E(1 - q_E)S_E(q_E, p(\bar{\beta}(q_E, q_V), q_E, q_V)), \\ \dot{q}_V &= q_V(1 - q_V)S_V(q_V, p(\bar{\beta}(q_E, q_V), q_E, q_V)). \end{aligned} \quad (10)$$

This system (10) is a Kolmogorov prey-predator dynamic system when

$$p(\bar{\beta}(q_{\mathrm{E}}, q_{\mathrm{V}}), q_{\mathrm{E}}, q_{\mathrm{V}}) = \hat{p}(q_{\mathrm{E}}, q_{\mathrm{V}})$$

is decreasing in q_{V} and increasing in q_{E}. Cultural type E_1 is the prey, and cultural type V_1 is the predator. It can be shown that under specific conditions, such a type of system may admit sustained oscillations and limit cycles (Kolmogorov 1936). To be more precise, denote

$$\Gamma_{\mathrm{E}} = \left(\frac{\partial p}{\partial \beta} \cdot \frac{(\partial p^{\mathrm{com}}/\partial q_{\mathrm{E}}) - (\partial p/\partial q_{\mathrm{E}})}{(\partial p/\partial \beta) - (\partial p^{\mathrm{com}}/\partial \beta)} + \frac{\partial p}{\partial q_{\mathrm{E}}}\right) \frac{\partial S_{\mathrm{E}}/\partial p}{-(\partial S_{\mathrm{E}}/\partial q_{\mathrm{E}})} > 0,$$

$$\Gamma_{\mathrm{V}} = \left(\frac{\partial p}{\partial \beta} \cdot \frac{(\partial p^{\mathrm{com}}/\partial q_{\mathrm{V}}) - (\partial p/\partial q_{\mathrm{V}})}{(\partial p/\partial \beta) - (\partial p^{\mathrm{com}}/\partial \beta)} + \frac{\partial p}{\partial q_{\mathrm{V}}}\right) \frac{\partial S_{\mathrm{V}}/\partial p}{-(\partial S_{\mathrm{V}}/\partial q_{\mathrm{V}})} < 0,$$

noting that $\partial \hat{p}(q_{\mathrm{E}}, q_{\mathrm{V}})/\partial q_g = \Gamma_g((\partial S_g/\partial q_g)/-(\partial S_g/\partial p))$ for $g = \mathrm{E}, \mathrm{V}$.

The system (10) has a unique interior steady state when (9) is satisfied. Noting that $\bar{q}_g'(p) = -(\partial S_g/\partial p)/(\partial S_g/\partial q_g)$, we rewrite equation (9) as $1 - \Gamma_1 - \Gamma_2 > 0$.

The linearized version of the system (10) at this interior steady state is given by

$$\begin{pmatrix} \dot{q}_{\mathrm{E}} \\ \dot{q}_{\mathrm{V}} \end{pmatrix} = \begin{pmatrix} q_{\mathrm{E}}^*(1 - q_{\mathrm{E}}^*)\dfrac{\partial S_{\mathrm{E}}}{\partial q_{\mathrm{E}}}[1 - \Gamma_{\mathrm{E}}] & q_{\mathrm{E}}^*(1 - q_{\mathrm{E}}^*)\left[\Gamma_{\mathrm{V}}\dfrac{\partial S_{\mathrm{E}}}{\partial p}\dfrac{-(\partial S_{\mathrm{V}}/\partial q_{\mathrm{V}})}{\partial S_{\mathrm{V}}/\partial p}\right] \\ q_{\mathrm{V}}^*(1 - q_{\mathrm{V}}^*)\left[\Gamma_{\mathrm{E}}\dfrac{\partial S_{\mathrm{V}}}{\partial p}\dfrac{-(\partial S_{\mathrm{E}}/\partial q_{\mathrm{E}})}{\partial S_{\mathrm{E}}/\partial p}\right] & q_{\mathrm{V}}^*(1 - q_{\mathrm{V}}^*)\dfrac{\partial S_{\mathrm{V}}}{\partial q_{\mathrm{V}}}[1 - \Gamma_{\mathrm{V}}] \end{pmatrix} \begin{pmatrix} q_{\mathrm{E}} \\ q_{\mathrm{V}} \end{pmatrix}.$$

The determinant Δ is

$$\Delta = q_{\mathrm{E}}^*(1 - q_{\mathrm{E}}^*)q_{\mathrm{V}}^*(1 - q_{\mathrm{V}}^*)\frac{\partial S_{\mathrm{E}}}{\partial q_{\mathrm{E}}}\frac{\partial S_{\mathrm{V}}}{\partial q_{\mathrm{V}}}[1 - \Gamma_{\mathrm{E}} - \Gamma_{\mathrm{V}}] > 0,$$

and the sufficient condition (9) for a unique interior steady state implies that $\Delta > 0$.

Now the trace T is given by

$$T = \frac{\partial S_{\mathrm{E}}}{\partial q_{\mathrm{E}}} q_{\mathrm{E}}^*(1 - q_{\mathrm{E}}^*)(1 - \Gamma_{\mathrm{E}}) + \frac{\partial S_{\mathrm{V}}}{\partial q_{\mathrm{V}}} q_{\mathrm{V}}^*(1 - q_{\mathrm{V}}^*)(1 - \Gamma_{\mathrm{V}}).$$

For the interior steady state to be unstable, it is necessary that $T > 0$ or that culture has to be sufficiently complement to institutions in one dimension E_1 (i.e., $\Gamma_{\mathrm{E}} > 0$) and sufficiently substitute in the other dimension V_1 (i.e., $\Gamma_{\mathrm{V}} < 0$). In such a case, a sufficient condition for both (9) and $T > 0$ is that Γ_{E} and Γ_{V} satisfy the following conditions:

$$1 - \Gamma_{\mathrm{V}} > \Gamma_{\mathrm{E}} > 1 + \frac{\partial S_{\mathrm{V}}/\partial q_{\mathrm{V}}}{\partial S_{\mathrm{E}}/\partial q_{\mathrm{E}}}\frac{q_{\mathrm{V}}^*(1 - q_{\mathrm{V}}^*)}{q_{\mathrm{E}}^*(1 - q_{\mathrm{E}}^*)}(1 - \Gamma_{\mathrm{V}}). \tag{11}$$

In that case, the unique interior steady state is unstable. The dynamic system stays inside the bounded region $[\epsilon, 1 - \epsilon]^2$ for ϵ small enough and such that $(q_{\mathrm{E}}^*, q_{\mathrm{V}}^*) \in [\epsilon, 1 - \epsilon]^2$. As a consequence of the Bendixon-Pointcaré theorem, there is then a limit cycle inside the region $[\epsilon, 1 - \epsilon]^2$.

Intuitively, an increase in the size q_E of the prey cultural subgroup E_1 induces a (more salient) externality to be internalized through an increase in the policy p (hence with institutional change increasing the political power β of group E); on the contrary, an increase in the size q_V of the predator cultural subgroup V_1 induces an externality to be internalized through a decrease in the policy p (hence decreasing β). Under these conditions, an increase in q_E—the prey—induces an increase in p and hence in q_V—the predator—but the increase in q_V—the predator—induces a decrease in p, which in turn feeds on the fraction of the prey q_E. This is the predator-prey mechanism that admits sustained oscillations and eventually a limit cycle when (11) is satisfied.

Appendix B

Results on the Dynamical System

In this appendix, we study in some detail the dynamics of our economy. The general dynamics in the text is characterized by equations 5–8. In fact, in this appendix, we study the specification characterized by equations 7 and 8 to simplify the formal analysis with little loss of generality.

We impose the following assumptions:

Assumption 1. Utility functions are such that $\mathbf{e}(\beta_t, q_t)$, $\mathbf{e}^{\text{com}}(\beta_t, q_t)$ are continuous functions.

Assumption 2. For regularity, we assume that $p(\beta_t, q_t)$ is monotonic in β_t and that all maps $p(\beta, q)$, $S(\beta, q)$ are smooth.

We consider our conceptualization that institutions at time $t + 1$ are designed at time t as a solution to

$$\max_{\beta\in[0,1]} W(\beta_t; \mathbf{a}(\beta, q_{t+1}), p(\beta, q_{t+1}), q_{t+1}). \tag{12}$$

It is useful to denote $\tilde{\mathbf{a}}(p, q)$ the vector of actions that describes the Nash equilibrium actions of society for a given policy level p, that is,

$$\tilde{\mathbf{a}}(p, q) = \mathbf{a} \text{ such that } \begin{cases} a \in \arg\max_a u(\mathbf{a}, p, q), \\ a^i \in \arg\max_{a^i} u^i(\mathbf{a}, p, q),\ i = 1, 2. \end{cases}$$

Note that $\mathbf{a}^{\text{com}}(\beta, q) = \tilde{\mathbf{a}}(p^{\text{com}}(\beta, q), q)$, and $\mathbf{a}(\beta, q) = \tilde{\mathbf{a}}(p(\beta, q), q)$. Given this and the notations

$$\mathbf{e}(\beta, q) = \begin{pmatrix} \mathbf{a}(\beta, q) \\ p(\beta, q) \end{pmatrix}$$

and

$$\mathbf{e}^{\text{com}}(\beta, q) = \begin{pmatrix} \mathbf{a}^{\text{com}}(\beta, q) \\ p^{\text{com}}(\beta, q) \end{pmatrix},$$

one has

$$p^{\text{com}}(\beta, q) = \arg\max_p W(\beta; \tilde{\mathbf{a}}(p, q), p, q)$$

and

$$\max_{\beta\in[0,1]} W(\beta_t; \mathbf{a}(\beta, q_{t+1}), p(\beta, q_{t+1}), q_{t+1}) = \max_{\beta} W(\beta_t; \mathbf{e}(\beta, q_{t+1}), q_{t+1}).$$

It is then easy to see that

$$\begin{aligned} W(\beta_t; e^{\text{com}}(\beta_t, q_{t+1}), q_{t+1}) &= W(\beta_t; \tilde{\mathbf{a}}(p^{\text{com}}(\beta, q_{t+1})), q_{t+1}), p^{\text{com}}(\beta, q_{t+1}), q_{t+1}) \\ &\geq W(\beta_t; \tilde{\mathbf{a}}(p, q_{t+1}), q_{t+1}), p, q_{t+1}) \text{ for all } p. \end{aligned}$$

In particular, for $p = p(\beta, q_{t+1})$ for all $\beta \in [0, 1]$, $\tilde{\mathbf{a}}(p(\beta, q_{t+1}), q_{t+1}) = \mathbf{a}(\beta, q_{t+1})$ and

$$\begin{aligned} W(\beta_t; e^{\text{com}}(\beta_t, q_{t+1}), q_{t+1}) &\geq W(\beta_t; \mathbf{a}(\beta, q_{t+1}), p(\beta, q_{t+1}), q_{t+1}) \\ &= W(\beta_t; \mathbf{e}(\beta, q_{t+1}), q_{t+1}). \end{aligned}$$

Thus, for all $\beta \in [0, 1]$, we have $W(\beta_t; e^{\text{com}}(\beta_t, q_{t+1}), q_{t+1}) \geq \max_\beta W(\beta_t; \mathbf{e}(\beta, q_{t+1}), q_{t+1})$. Now, suppose that there is a value $\beta \in [0, 1]$ such that $p^{\text{com}}(\beta_t, q_{t+1}) = p(\beta, q_{t+1})$; then at such β, one has $\mathbf{e}(\beta, q_{t+1}) = e^{\text{com}}(\beta_t, q_{t+1})$ and consequently $\beta = \arg\max_{\beta'} W(\beta_t; \mathbf{e}(\beta', q_{t+1}), q_{t+1})$.

Suppose alternatively that for all $\beta \in [0, 1] p^{\text{com}}(\beta_t, q_{t+1}) - p(\beta, q_{t+1}) > 0$ at any β. Then the assumption that $p(\beta, q_{t+1})$ is monotonic implies that $p(\beta, q_{t+1})$ is the closest to $p^{\text{com}}(\beta_t, q_{t+1})$ at $\beta = 0$ or at 1, depending on the sign of variation of $p(\beta, q_{t+1})$ with respect to β. In that case, $\beta = 0$ or $\beta = 1$ is the solution of the maximization problem (12). A similar argument follows for the case where for all $\beta \in [0, 1] p^{\text{com}}(\beta_t, q_{t+1}) - p(\beta, q_{t+1}) < 0$ at any β.

We then study the dynamics of $(\beta_t, q_t) \in [0, 1]^2$. The fundamental institutional and cultural dynamics can be conveniently rewritten with the map $f : [0, 1]^2 \to [0, 1]$ as follows:

$$\beta_{t+1} = f(\beta_t, q_{t+1}) := \begin{cases} \beta \text{ such that } p^{\text{com}}(\beta_t, q_{t+1}) = p(\beta, q_{t+1}) & \text{if it exists,} \\ \left[\begin{array}{l} 1 \ \text{ if } P(\beta_t, q_{t+1}) > 0, \forall\ 0 \leq \beta \leq 1, \\ 0 \ \text{ if } P(\beta_t, , q_{t+1}) < 0, \forall\ 0 \leq \beta \leq 1, \end{array}\right. & \text{else} \end{cases} \tag{13}$$

$$q_{t+1} - q_t := q_t(1 - q_t)S(\beta_{t+1}, q_{t+1}).$$

We shall study the dynamical system in the continuous time limit, where the change in β_t and q_t between time t and $t + dt$ is, respectively, λdt and μdt for $dt \to 0$. Denoting $g : [0, 1]^2 \to \mathbb{R}$ such that $g(\beta, q) = q(1 - q)S(\beta, q)$, we get[69]

[69] As is well known, discrete time dynamics may generate complex dynamic behaviors that are difficult to characterize and go beyond the points we want to emphasize about the coevolution between culture and institutions.

System (14) is easily obtained in the following way. We assume that between time t and $t + dt$, an opportunity to change institutions arises with an instantaneous rate λdt. Therefore, the dynamics of β is written as

$$\beta_{t+dt} = (1 - \lambda dt)\beta_t + \lambda dt f(\beta_t, q_{t+dt}).$$

Similarly, we may assume that between t and $t + dt$, a fraction μdt of individuals just before dying have an offspring socialized through cultural transmission. Then, the dynamics of q is written as

$$q_{t+dt} = (1 - \mu dt)q_t + \mu dt[q_t + q_t(1 - q_t)S(\beta_{t+dt}, q_{t+dt})].$$

Letting $dt \to 0$ provides immediately $\dot{\beta} = \lambda[f(\beta, q) - \beta]$ and $\dot{q} = \mu q(1 - q)S(\beta, q) = \mu g(\beta, q)$.

$$\begin{aligned} \dot{\beta} &= \lambda[f(\beta, q) - \beta], \\ \dot{q} &= \mu q(1-q)S(\beta, q) = \mu g(\beta, q), \end{aligned} \tag{14}$$

given the initial conditions (β_0, q_0).

B1. *The Dynamics of* β *Given* q

LEMMA B1. Under assumptions 1 and 2, $f:[0,1]^2 \to [0,1]$ is a continuous function in $(\beta, q) \in [0,1]^2$.

Proof. First of all, note that when $p(\beta_{t+1}, q_{t+1}) = p^{\mathrm{com}}(\beta_t, q_{t+1})$ is not satisfied for any β_{t+1}, for some q_{t+1}, the assumption that $p(\beta, q)$ is monotonic implies that β_{t+1} is equal to 0 or 1, depending on the sign of $p^{\mathrm{com}}(\beta_{t+1}, q_{t+1}) - p(\beta_t, q_{t+1})$. In the continuous time limit, $q_{t+1} = q_t = q$, and hence in this case, trivially, f maps continuously $(\beta, q) \in [0,1]^2$ into $\{0\}$.

Consider equation (13) again. We show that β_{t+1} is a continuous function of β_t and q_{t+1} when $p(\beta_{t+1}, q_{t+1}) = p^{\mathrm{com}}(\beta_t, q_{t+1})$ is satisfied. To this end, note that the assumed monotonicity in β of $p(\beta, q)$ implies that when $p(\beta_{t+1}, q_{t+1}) = p^{\mathrm{com}}(\beta_t, q_{t+1})$ is satisfied, we can write $\beta_{t+1} = p^{-1}(p, q_{t+1})$ and hence $\beta_{t+1} = p^{-1}(p^{\mathrm{com}}(\beta_t, q_{t+1}), q_{t+1})$, a continuous function. Again, in the continuous time limit, $q_{t+1} = q_t$, and hence we can construct a continuous function $f:[0,1]^2 \to \mathbb{R}$ such that $\dot{\beta}_t = f(\beta_t, q_t) - \beta$.

Finally, it is straightforward to see that as $p^{\mathrm{com}}(\beta_{t+1}, q_{t+1}) - p(\beta_t, q_{t+1})$ crosses 0, $\beta_{t+1} = p^{-1}(p^{\mathrm{com}}(\beta_t, q_{t+1}), q_{t+1})$ converges continuously to 0 or 1 depending on the direction of the crossing so as to preserve continuity. QED

The stationary states of the dynamics of β satisfy $\beta = f(\beta, q)$, which is equivalent to β being a zero of $P(\beta, q) := p^{\mathrm{com}}(\beta, q) - p(\beta, q)$. Let then $\pi:[0,1] \to [0,1]$ maps q into the stationary states β such that $P(\beta, q) = 0$; that is, the map π satisfies $P(\pi(q), q) = 0$. We consider only the regular case in which $P(\beta, q) \neq 0$ at the vertices of $[0,1]^2$, leaving the simple but tedious analysis of the singular cases to the reader. Also, we say that q is a regular point of $\beta \in \pi(q)$ if any stationary state $\beta \in \pi(q)$ satisfies that property that $(\partial P(\beta, q)/\partial \beta) \neq 0$, that is, if $p(\beta, q)$ and $p^{\mathrm{com}}(\beta, q)$ intersect transversally.

LEMMA B2. Under assumptions 1 and 2, the map $\pi:[0,1] \to [0,1]$ is a nonempty and compact valued upper hemicontinuous correspondence with connected components.

Proof. The proof is a direct consequence of the continuity of f proved in lemma B1 and the fact that $\beta \in \pi(q)$ is equivalent to $\beta = f(\beta, q)$. QED

PROPOSITION B1. Under assumptions 1 and 2, the dynamics of β as a function of $q \in [0,1]$ has the following properties:

1. $P(0, q) > 0$, $P(1, q) < 0$ for any $q \in [0,1]$, and $p(\beta, q)$ is increasing; or $P(0, q) < 0$, $P(1, q) > 0$ for any $q \in [0,1]$, and $p(\beta, q)$ is decreasing. For any given regular $q \in [0,1]$, there exists an odd number of regular stationary states $\beta \in \pi(q)$; furthermore $\beta = 0, 1$ are also stationary states for given $q \in [0,1]$. The stability properties of the regular stationary states alternate

with the smallest and the larger being always locally stable; the boundaries $\beta = 0, 1$ are locally unstable for all $q \in [0, 1]$.

2. $P(0, q) < 0$, $P(1, q) > 0$ for any $q \in [0, 1]$, and $p(\beta, q)$ is increasing; or $P(0, q) > 0$, $P(1, q) < 0$ for any $q \in [0, 1]$, and $p(\beta, q)$ is decreasing. For any given $q \in [0, 1]$, there exists an odd number of regular stationary states $\beta \in \pi(q)$; furthermore, $\beta = 0, 1$ are also stationary states for given $q \in [0, 1]$. The stability properties of the regular stationary states alternate with the smallest and the larger being always locally unstable; the boundaries $\beta = 0, 1$ are locally stable.
3. $P(0, q) < 0$, $P(1, q) < 0$ for any $q \in [0, 1]$. For any given $q \in [0, 1]$, there exists either none or an even number of regular stationary states $\beta \in \pi(q)$; furthermore, $\beta = 0$ is also a stationary state for given $q \in [0, 1]$. The stability properties of the regular stationary states alternate with the smallest always locally unstable; the boundary $\beta = 0$ is locally stable.
4. $P(0, q) > 0$, $P(1, q) > 0$ for any $q \in [0, 1]$. For any given $q \in [0, 1]$, there exists either none or an even number of regular stationary states $\beta \in \pi(q)$; furthermore, $\beta = 1$ is also a stationary state for given $q \in [0, 1]$. The stability properties of the regular stationary states alternate with the smallest always locally stable; the boundary $\beta = 1$ is locally stable.
5. $P(0, q)$ and/or $P(1, q)$ change sign with $q \in [0, 1]$. The characterization obtained above then can be repeated for each subinterval of [0, 1] in which the Brouwer degree of the manifold $\pi(q)$ is invariant. We leave the tedious categorization of all possible cases to the reader.

B2. *The Dynamics of* q *Given* β

LEMMA B3. Under assumptions 1 and 2, $g : [0, 1]^2 \to [0, 1]$ such that $g(q) = q(1 - q)S(\beta, q)$ is a continuous function in $(\beta, q) \in [0, 1]^2$.

Proof. The proof is an immediate consequence of the continuity of $S(\beta, q)$ that are associated with assumptions 1 and 2. QED

Let $\sigma : [0, 1] \to [0, 1]$ map β into the zeros of $S(\beta, q)$, that is, the cultural states q such that $S(\beta, q) = 0$; the map σ satisfies $S(\sigma(\beta), \beta) = 0$. Consider the regular case in which $S(\beta, q) \neq 0$ at the vertices of $[0, 1]^2$, leaving the simple but tedious analysis of the singular cases to the reader. We say that β is a regular point of $q \in \sigma(\beta)$ if any stationary state $q \in \sigma(\beta)$ satisfies that property that $(\partial S(\beta, q)/\partial q) \neq 0$, that is, if $d^i(\beta, q)$ and $d^j(1 - \beta, 1 - q)$ intersect transversally. The characterization of the stationary states of the cultural dynamics at given institutions β are obtained when $g(\beta, q) = 0$ and depend crucially on the topological properties of the zeros of $S(\beta, q)$.

LEMMA B4. Under assumptions 1 and 2, the stationary states of the cultural dynamics are characterized by the map $\sigma : [0, 1] \to [0, 1]$, which is a nonempty and compact valued upper hemicontinuous correspondence with connected components. Moreover, $q = 0$ and $q = 1$ are also stationary states for any $0 \leq \beta \leq 1$.

Proof. The proof is a direct consequence of the continuity of g, proved in lemma A1, and of the fact that $g(0, \beta) = g(1, \beta) = 0$ for any $0 \leq \beta \leq 1$. QED

PROPOSITION B2. Under assumptions 1 and 2, the dynamics of q as a function of $\beta \in [0, 1]$ has the following properties:

1. $S(\beta, 0) < 0,\ S(\beta, 1) > 0$ for any $\beta \in [0, 1]$. For any given regular $\beta \in [0, 1]$, there exists an odd number of regular stationary states $q \in \sigma(\beta)$. By lemma B4, $q = 0, 1$ are also stationary states for given $\beta \in [0, 1]$. The stability properties of the regular stationary states alternate, starting with the $q = 0$ being stable and ending with $q = 1$ also being stable. If the dynamics supports a unique interior stationary state q^*, then it is unstable.
2. $S(\beta, 0) > 0,\ S(\beta, 1) < 0$ for any $q \in [0, 1]$. For any given regular $\beta \in [0, 1]$, there exists an odd number of regular stationary states $q \in \sigma(\beta)$. By lemma B4, $q = 0, 1$ are also stationary states for given $\beta \in [0, 1]$. The stability properties of the regular stationary states alternate, starting with the $q = 0$ being unstable and ending with $q = 1$ also being unstable. If the dynamics supports a unique interior stationary state q^*, then it is stable.
3. $S(\beta, 0) < 0,\ S(\beta, 1) < 0$ for any $\beta \in [0, 1]$. For any given regular $\beta \in [0, 1]$, there exists either none or an even number of regular stationary states $q \in \sigma(\beta)$. By lemma B4, $q = 0, 1$ are also stationary states for given $\beta \in [0, 1]$. The stability properties of the regular stationary states alternate, starting with the $q = 0$ being stable and ending with $q = 1$ being unstable.
4. $S(\beta, 0) > 0,\ S(\beta, 1) > 0$ for any $\beta \in [0, 1]$. For any given regular $\beta \in [0, 1]$, there exists either none or an even number of regular stationary states $q \in \sigma(\beta)$. By lemma B4, $q = 0, 1$ are also stationary states for given $\beta \in [0, 1]$. The stability properties of the regular stationary states alternate, starting with the $q = 0$ being unstable and ending with $q = 1$ being stable.
5. $S(\beta, 0)$ and/or $S(\beta, 1)$ change sign with $\beta \in [0, 1]$. The characterization obtained above then can be repeated for each subinterval of [0, 1] in which the Brouwer degree of the manifold $\sigma(\beta)$ is invariant. We leave the tedious categorization of all possible cases to the reader.

Proof. Under assumptions 1 and 2, $S(\beta, q)$ is smooth and (β, q) lie in the compact set $[0, 1]^2$. $\sigma(\beta)$ is a one-dimensional smooth manifold with boundary by a general version of the implicit function theorem (see, e.g., Milnor [1965] 1997, lemma 4, 13). The statement is then proved closely along the lines of the proof of proposition B1, using the full characterization of one-dimensional manifolds and Brouwer degree theory, thinking of $S(\beta, q)$ as a homothopy function varying β. We leave the details to the reader. QED

B3. The Joint Dynamics of (β, q)

The dynamical system (14), even under assumptions 1 and 2, is impossible to study in general. We can, however, show that at least one stationary state always exists and characterize sufficient conditions for the existence of an interior stationary state. To this end, we restate here more formally proposition 5.

Proposition B3. Under assumptions 1 and 2, the dynamical system ((5), (8)) has at least one stationary state. Furthermore, if the Brouwer degree of both $\pi(q)$ and $\sigma(\beta)$ is ± 1, the dynamical system has at least one interior stationary state.

Proof. The proof of the existence of a stationary state is a direct consequence of the characterization of $\pi(q)$ and $\sigma(\beta)$ in lemmas B2 and B4.

The proof of the existence of an interior stationary state under the Brouwer degree conditions is a consequence of the Jordan curve theorem, which we state in the following for completeness:[70]

A curve J in $\mathbb{R}^2$, which is the image of an injective continuous map of a circle into $\mathbb{R}^2$, has two components (an inside and outside), with J the boundary of each.

Figure B1 represents a Jordan curve J on the plane.

Consider the compact space $[0, 1]^2 \subset \mathbb{R}^2$ in which (β, q) lies. By lemma 4, the locus of stationary states of the cultural dynamics contains the boundaries $q = 0, 1$ as well as the map $\sigma(\beta)$, which—in the case its Brouwer degree is ± 1—is homeomorphic to the compact interval $[0, 1]$. The map $\pi(q)$ is also homeomorphic to the compact interval $[0, 1]$ in the case its Brouwer degree is ± 1.

We can therefore construct a Jordan curve J composed of $\pi(q)$, $(\beta > \pi(0), q = 0)$, $(\beta > \pi(1), q = 1)$, $\beta = 1$. Since $\sigma(\beta)$ connects the $\beta = 1$ and $\beta = 0$, it has a component inside and one outside the curve J. Furthermore, $0 < \sigma(\beta) < 1$ by construction. The Jordan curve theorem then guarantees that $\pi(q)$ and $\sigma(\beta)$ cross in the interior of $[0, 1]^2$; see figure B2 for a graphical representation of the construction. QED

Note that propositions B1 and B2 provide conditions, respectively on $p(\beta, q)$ and $S(\beta, q)$, guaranteeing that the Brouwer degree of $\pi(q)$ and $\sigma(\beta)$ is ± 1. Also, the analysis leading to proposition B3 can be extended to dynamical systems in which the Brower degrees of $\pi(q)$ and $\sigma(\beta)$ are not invariant.

B4. Further Characterization of the Joint Dynamics

We provide here the proof of propositions and results in the text that characterize the complex dynamics of culture and institutions.

B4.1. Complementarity and Substitution between Institutions and Culture, Cycles, and Oscillations

Proof of proposition 5. Suppose that conditions (15) are satisfied at an interior steady state (β^*, q^*) of the system (14):

$$\frac{\partial P(\beta^*, q^*)}{\partial \beta}, \frac{\partial S(\beta^*, q^*)}{\partial q} < 0. \tag{15}$$

The linearized local dynamics around the interior steady state (β^*, q^*) can then easily be obtained by

$$\begin{pmatrix} \dot{\beta} \\ \dot{q} \end{pmatrix} = \begin{pmatrix} \lambda \left[\frac{\partial P/\partial \beta}{p_\beta}\right]_{(\beta^*, q^*)} & \lambda \left[\frac{\partial P/\partial q}{p_\beta}\right]_{(\beta^*, q^*)} \\ \mu q^*(1 - q^*)\left[\frac{\partial S}{\partial \beta}\right]_{(\beta^*, q^*)} & \mu q^*(1 - q^*)\left[\frac{\partial S}{\partial q}\right]_{(\beta^*, q^*)} \end{pmatrix} \begin{pmatrix} \beta \\ q \end{pmatrix}. \tag{16}$$

[70] The theorem is a standard result in algebraic topology; see Hatcher (2002, 169) for a proof.

If we recall (1) and (2) and our normalization $p_\beta > 0$, then with enough regularity of the policy functions p^{com} and p, there exists a connected neighborhood of (β^*, q^*) such that the trace

$$T = \lambda\left[\frac{\partial P/\partial\beta}{p_\beta}\right] + \mu q(1-q)\left[\frac{\partial S}{\partial q}\right]$$

is negative and therefore does not change sign on that domain. In this case, the Bendixson negative criterion then precludes the existence of local periodic orbits or limit cycles around (β^*, q^*) in that domain.

Note that when (15) are globally satisfied for all $(\beta, q) \in [0,1]\times[0,1]$, it is not possible to get globally periodic orbits and limit cycles for dynamical system (14). Indeed, given that in the simple connected domain $D = [0,1]\times[0,1]$ the sign of the trace

$$T = \lambda\left[\frac{\partial P/\partial\beta}{p_\beta}\right] + \mu q(1-q)\left[\frac{\partial S}{\partial q}\right]$$

is always strictly negative, the Bendixson negative criterion again precludes the existence of periodic orbits of (14) in this domain. QED

Proof of proposition 6. Consider first an interior steady state (β^*, q^*) of (14) that is locally stable. Given our normalization $p_\beta > 0$ and the linearized system (16), we have the standard Hessian conditions

$$\frac{\partial P(\beta^*, q^*)}{\partial\beta} < 0, \frac{\partial S(\beta^*, q^*)}{\partial q} < 0,$$
$$\left[\frac{\partial P}{\partial\beta}\cdot\frac{\partial S}{\partial q} - \frac{\partial P}{\partial q}\cdot\frac{\partial S}{\partial\beta}\right]_{(\beta^*, q^*)} > 0.$$

These local stability conditions ensure that the trace $T < 0$ and that the determinant $\Delta > 0$. Dampened oscillations (a stable spiral steady state equilibrium) require $T^2 < 4\Delta$. This last condition is written as

$$\left[\lambda\left[\frac{\partial P/\partial\beta}{p_\beta}\right]_{(\beta^*, q^*)} + \mu q^*(1-q^*)\left[\frac{\partial S}{\partial q}\right]_{(\beta^*, q^*)}\right]^2$$
$$< \frac{4\lambda\mu q^*(1-q^*)}{p_\beta(\beta^*, q^*)}\left[\frac{\partial P}{\partial\beta}\cdot\frac{\partial S}{\partial q} - \frac{\partial P}{\partial q}\cdot\frac{\partial S}{\partial\beta}\right]_{(\beta^*, q^*)},$$

or, after manipulations,

$$\left[\lambda\left[\frac{\partial P/\partial\beta}{p_\beta}\right]_{(\beta^*, q^*)} - \mu q^*(1-q^*)\left[\frac{\partial S}{\partial q}\right]_{(\beta^*, q^*)}\right]^2 < -\frac{4\lambda\mu q^*(1-q^*)}{p_\beta(\beta^*, q^*)}\left[\frac{\partial P}{\partial q}\cdot\frac{\partial S}{\partial\beta}\right]_{(\beta^*, q^*)}. \quad (17)$$

Given our normalization $p_\beta > 0$, when institutions and culture are dynamic complements at (β^*, q^*), $(\partial P(\beta^*, q^*))/\partial q$ and $(\partial S(\beta^*, q^*))/\partial\beta$ have the same sign. Hence, $[(\partial P/\partial q)\cdot(\partial S/\partial\beta)] > 0$, and the right-hand side of inequality (17) is negative. Given that the left-hand side is positive, it follows that (17) cannot be satisfied and there are no dampening oscillations in cultural and institutional change when institutions and culture are dynamic complements at (β^*, q^*). QED

Proof of the existence of dampened oscillations when culture and institutions are dynamic substitutes. Assume now that culture and institutions are dynamic substitutes at the interior locally stable steady state (β^*, q^*). This implies that

$$\left[\frac{\partial P}{\partial q}\cdot\frac{\partial S}{\partial \beta}\right]_{(\beta^*,q^*)} < 0,$$

and dampening oscillations occur when (17) is satisfied. In this case, nonmonotonic dynamics in culture and institutions obtain when

$$\frac{\lambda}{\mu}\left[\left[\frac{\partial P/\partial\beta}{p_\beta}\right]_{(\beta^*,q^*)} - \frac{\mu}{\lambda}q^*(1-q^*)\left[\frac{\partial S}{\partial q}\right]_{(\beta^*,q^*)}\right]^2 < \frac{4q^*(1-q^*)}{p_\beta(\beta^*,q^*)}\left[-\frac{\partial P}{\partial q}\cdot\frac{\partial S}{\partial\beta}\right]_{(\beta^*,q^*)}. \quad (18)$$

Using the local stability conditions for the Hessian at (β^*, q^*), we have

$$\begin{aligned}\left[\frac{\partial P/\partial\beta}{p_\beta}\right]_{(\beta^*,q^*)} &= -a < 0,\\ -q^*(1-q^*)\left[\frac{\partial S}{\partial q}\right]_{(\beta^*,q^*)} &= b > 0,\\ \frac{4q^*(1-q^*)}{p_\beta(\beta^*,q^*)}\left[-\frac{\partial P}{\partial q}\cdot\frac{\partial S}{\partial\beta}\right]_{(\beta^*,q^*)} &= M > 0.\end{aligned}$$

Denoting $x = \mu/\lambda$ the relative rate of change between culture and institutions, we can write condition (18) as

$$(-a + bx)^2 < Mx\,. \quad (19)$$

Simple examination of this condition reveals that (19) is satisfied when $x \in (x_-; x_+)$, with

$$x_\pm = \frac{(2ab + M) \pm \sqrt{(2ab+M)^2 - 4(ab)^2}}{2b^2} > 0\,.$$

As a consequence, nonmonotonic dynamics of institutions and culture around the locally stable steady state (β^*, q^*) are obtained when institutions and culture are dynamic substitutes and the relative rate of change between culture and institutions is neither too low nor too high. QED

Proof of proposition 7. Consider the cultural multiplier m_{SS} at a locally stable interior steady state (β^*, q^*). Recall the required normalizations

$$p_\beta = \frac{\partial p(\beta^*, q^*, \gamma)}{\partial\beta} > 0,\ p_\gamma^{com} - p_\gamma = P_\gamma = \frac{\partial P(\beta^*, q^*, \gamma)}{\partial\gamma} > 0. \quad (20)$$

The comparative statics on (β^*, q^*) on the parameter are then easily obtained by differentiation of

$$\begin{aligned}P(\beta^*, q^*, \gamma) &= 0,\\ S(\beta^*, q^*) &= 0,\end{aligned} \quad (21)$$

and one gets

$$\frac{d\beta^*}{d\gamma} = \frac{-(\partial S/\partial q)P_\gamma}{(\partial P/\partial\beta)(\partial S/\partial q) - (\partial P/\partial q)(\partial S/\partial\beta)},$$
$$\frac{dq^*}{d\gamma} = \frac{(\partial S/\partial\beta)P_\gamma}{(\partial P/\partial\beta)(\partial S/\partial q) - (\partial P/\partial q)(\partial S/\partial\beta)}. \tag{22}$$

Consider now the impact of a change in γ on institutional change, fixing q to its preshock value. If we differentiate the first equation in (21),

$$\left(\frac{d\beta^*}{d\gamma}\right)_{q=q^*} = \frac{P_\gamma}{-(\partial P/\partial\beta)} > 0\,.$$

the stability condition for (β^*, q^*) requires $\partial P/\partial\beta < 0$, $\partial S/\partial q < 0$, and $\Delta = (\partial P/\partial\beta)(\partial S/\partial q) - (\partial P/\partial q)(\partial S/\partial\beta) > 0$. Coupled with condition (20), this implies that the cultural multiplier on institutional change m at (β^*, q^*), $m = (d\beta^*/d\gamma)/(d\beta^*/d\gamma)_{q=q^*} - 1$, is positive if and only if

$$\frac{\partial P}{\partial q}\cdot\frac{\partial S}{\partial\beta} > 0,$$

which is the condition for complementarity of the institutional and cultural dynamics. QED

Proof of proposition 8. Consider the cultural multiplier m_{DD} on institutional change from initial condition (β_0, q_0) in the basin of attraction B of a stationary state (β^*, q^*). In this case, the full dynamics of culture and institutions from (β_0, q_0) converges by construction to (β^*, q^*). In particular, institutions converge to $\beta^* = \beta(q^*)$ and $P(\beta(q^*), q^*) = 0$. In the counterfactual case in which the cultural composition of society had remained fixed, the dynamics of institutions would have converged to $\beta(q_0)$ and $P(\beta(q_0), q_0) = 0$. The cultural multiplier m_{DD} is written as

$$m_{\mathrm{DD}} = \frac{\beta(q^*)}{\beta(q_0)} - 1\,.$$

Assume first that $q_0 < q^*$. Because institutions and culture are global dynamic complements, we know from proposition 6 that there are no dampened oscillations, and therefore q_t monotonically increases from q_0 to q^*. Moreover, for all (β, q) in the basin of attraction B of (β^*, q^*), we should have $P_\beta(\beta, q) < 0$, as β converges to the stable manifold $\beta(q)$ for q inside the projection of the basin of attraction B on the space of $q \in [0, 1]$.

If we then first consider the case where $\partial P/\partial q > 0$, we get

$$P(\beta(q^*), q_0) < P(\beta(q^*), q^*) = 0 = P(\beta(q_0), q_0),$$

and we immediately obtain that $\beta^* = \beta(q^*) > \beta(q_0)$ and $m_{\mathrm{DD}} > 0$.

Similarly, when $\partial P/\partial q < 0$,

$$P(\beta(q^*), q_0) > P(\beta(q^*), q^*) = 0 = P(\beta(q_0), q_0),$$

and immediately $\beta^* = \beta(q^*) < \beta(q_0)$ and $m_{\mathrm{DD}} < 0$. Consequently, whatever the sign of $\partial P/\partial q$ is, m_{DD} has the same sign as $\partial P/\partial q$.

The case $q_0 > q^*$ can be handled by a similar argument, and one can easily see that m_{DD} should have the opposite sign as $\partial P/\partial q$ in such a case.

In conclusion, m_{DD} has the same sign as $[q^* - q_0] \cdot (\partial P/\partial q)$. QED

B4.2. Decomposition of the Cultural Multiplier on an Aggregate Variable $A(p, q, a^1(p), a^2(p))$

The cultural multiplier governs the effects of the interaction between culture and institutions on any aggregate economic variable of interest, for example, per capita income, public good provision, or any other measure of economic activity. Let $A(p, q, a(p), a^1(p), a^2(p))$ formally denote the economic aggregate. A *cultural multiplier on A* can then be defined as

$$m_A = \frac{dA}{d\gamma} \Big/ \left(\frac{dA}{d\gamma}\right)_{q=q^*} - 1.$$

Noting from (22) that

$$\frac{dq^*}{d\gamma} = \frac{d\beta^*}{d\gamma} \cdot \frac{\partial S/\partial \beta}{-(\partial S/\partial q)},$$

we get

$$\frac{dA}{d\gamma} = \Bigg\{ \underbrace{\left[A_p + \left(A_a a_p + A_{a^1} a_p^1 + A_{a^2} a_p^2\right)\right] p_\beta}_{\text{direct effect}} + \underbrace{\left[A_q + \left[A_p + \left(A_a a_p + A_{a^1} a_p^1 + A_{a^2} a_p^2\right)\right] p_q\right] \frac{\partial S/\partial \beta}{-(\partial S/\partial q)}}_{\text{indirect effect}} \Bigg\} \frac{d\beta^*}{d\gamma}.$$

The effect of γ on institutions will come from a direct effect as well as an indirect one. The direct effect in turn will be composed of two terms: a direct effect of the policy change induced by an institutional change p_β on the aggregate variable A (i.e., the term A_p in the first square brackets) and the impact of changes in private actions $a(p)$, $a^1(p)$ and $a^2(p)$ as induced also by the policy change p_β, (i.e., the term $A_a a_p + A_{a^1} a_p^1 + A_{a^2} a_p^2$ in the first square brackets). The indirect effect of cultural evolution comes from the compositional effect of changing the sizes of the populations with different cultural traits (A_q) plus again the change in policy and private actions $[A_p + (A_a a_p + A_{a^1} a_p^1 + A_{a^2} a_p^2)]p_q$, which such a cultural compositional change induces.

Furthermore,

$$\left(\frac{dA}{d\gamma}\right)_{q=q^*} = \left[A_p + \left(A_a a_p + A_{a^1} a_p^1 + A_{a^2} a_p^2\right)\right] p_\beta \cdot \left(\frac{d\beta^*}{d\gamma}\right)_{q=q^*}.$$

Recalling that the cultural multiplier on institutions is $m = [(d\beta^*/d\gamma)/(d\beta^*/d\gamma)_{q=q^*} - 1]$, we have

$$m_A = \frac{dA}{d\gamma} / \left(\frac{dA}{d\gamma}\right)_{q=q^*} - 1 = m + K \cdot (1 + m),$$

with

$$K = \frac{A_q + \left[A_p + \left(A_a a_p + A_{a^1} a_p^1 + A_{a^2} a_p^2\right)\right] p_q}{\left[A_p + \left(A_a a_p + A_{a^1} a_p^1 + A_{a^2} a_p^2\right)\right] p_\beta} \frac{\partial S / \partial \beta}{-(\partial S / \partial q)},$$

and hence m_A can be expressed as a function of the institutional cultural multiplier m in terms of two components.

The first term is the cultural multiplier itself, m. This reflects the direct pass-through multiplier effect of institutions β on the aggregate variable $A(.)$ (through the impact of β on the equilibrium policy p and individual behaviors). The second term $K \cdot (1 + m)$ represents another multiplier effect on $A(.)$ that is triggered by the impact of institutional change on the cultural dynamics q, which in turn also affect the aggregate variable $A(.)$ through population effects. More precisely, this term is proportional to the relative degree of institutional change under joint evolution $(1 + m) = (d\beta^*/d\gamma)/(d\beta^*/d\gamma)_{q=q^*}$, with a coefficient of proportionality K that reflects the sensitivity of cultural dynamics to institutions, as well as how the aggregate variable $A(.)$ depends on cultural change through a cultural compositional effect and a culturally induced policy shift (i.e., the effect of q on $p(\beta, q, \gamma)$). Depending on the sign of K, this second effect may either magnify or mitigate the direct pass-through cultural multiplier of institutional change on the variable A.

B4.3. Cultural Dynamics in Bisin and Verdier (2001)

Cultural transmission is modeled as the result of direct vertical (parental) socialization and horizontal/oblique socialization in the heterogeneous political group at large. *Direct vertical* socialization to the parent's trait $i \in I = \{1, 2\}$ occurs with probability d^i. If a child from a family with trait i is not directly socialized, which occurs with probability $1 - d^i$, he/she is *horizontally/obliquely* socialized by picking the trait of a role model chosen randomly in the population inside the political group (i.e., he/she picks trait i with probability q^i and trait $i' \neq i$ with probability $q^{i'}$).

If we let P^{ii} ($P^{ii'}$) denote the probability that a child of a family with cultural trait $i \in I$ is socialized to trait i (i'), we obtain

$$P^{ii} = d^i + (1 - d^i) q^i, \; P^{ii'} = (1 - d^i) q^{i'}$$

Let $V^{ii'}(\beta, q)$ denote the utility to a cultural trait i parent of a type i' child. It depends on the institutional setup β and the cultural distribution $q = q^1 = 1 - q^2$ the child will face when he/she will make his/her economic decision $a^{i'}$.[71] Let $C(d^i)$ denote socialization costs. Direct socialization for any $i \in I = \{1, 2\}$ is then the solution to the following parental socialization problem:

[71] In extensive notation, $V^{ii'}(\beta, q) = u^i(a^{i'}(\beta, q), p(\beta, q); a(\beta, q), q)$.

$$\max_{d^i \in [0,1]} - C(d^i) + \sum_{i' \in I} P^{ii'} V^{ii'}(\beta, q)$$

$$\text{such that } P^{ii} = d^i + (1 - d^i)q^i, \ P^{ii'} = (1 - d^i)q^{i'}.$$

As usual in this literature, define $\Delta V^i(\beta, q) = V^{ii}(\beta, q) - V^{ii'}(\beta, q)$ as the *cultural intolerance* associated with trait i. It follows that direct socialization, $d^i(\beta, q)$, with some notational abuse is determined by the first-order conditions

$$C'(d^1) = (1 - q)\Delta V^1(\beta, q),$$
$$C'(d^2) = (1 - q)\Delta V^2(\beta, q).$$

When we turn again to the explicit notation for time t, the dynamics of q_t is straightforwardly determined by

$$q_{t+1} - q_t = q_t(1 - q_t)S(\beta_{t+1}, q_{t+1}),$$

with $S(\beta, q) = d^1(\beta, q) - d^2(\beta, q)$.

It is convenient to impose for regularity the following assumption (we do so in the examples) of separability of preference structures and quadratic costs of socialization:

Assumption 3. $u^i(a^i, p; \mathbf{a}, q) = v^i(a^i, p) + H(p; \mathbf{a}, q)$, and $C(d^i) = (1/2)(d^i)^2$ for type $i = 1, 2$.

Under (3), the cultural replicator dynamics (in continuous time) for fixed institutions β become

$$\dot{q_t} = q_t(1 - q_t)S(\beta, q_t),$$

with $S(\beta, q)$ rewritten as

$$S(\beta, q) = \left[\Delta V^1(p(\beta, q)) \cdot (1 - q)\right] - [\Delta V^2(p(\beta, q)) \cdot q)],$$

with

$$\Delta V^1(p) = v^1(a^1(p), p) - v^1(a^2(p), p),$$
$$\Delta V^2(p) = v^2(a^2(p), p) - v^2(a^1(p), p).$$

Any interior stationary state q^* is obtained as a solution to

$$\frac{\Delta V^1(p(\beta, q))}{\Delta V^2(p(\beta, q))} = \frac{q}{1 - q}. \tag{23}$$

This equation may have many solutions characterizing the cultural steady state manifold $q = q(\beta)$. One may, however, provide sufficient conditions for the existence of a unique convergent cultural steady state. Specifically:

i. Assume that $\Delta V^1(p(\beta, 0)) > 0$, $\Delta V^2(p(\beta, 1)) > 0$, and the function $\phi(\beta, q) = \log[\Delta V^1(p(\beta, q))/\Delta V^2(p(\beta, q)))]$ is such that $\phi'_q(\beta, q) < 4$ for all $(\beta, q) \in [0, 1]^2$. Then (23) defines a unique solution $q(\beta)$ for every value of β, and the cultural dynamics tend to that interior stationary state $q(\beta)$, whereby q increases when $q < q(\beta)$ and decreases instead when $q > q(\beta)$.

To see that, notice that at a point q^* satisfying $S(\beta, q^*) = 0$ or equivalently (23), one has

$$\frac{\partial S}{\partial q}\Big|_{q^*} = \left[\frac{d\Delta V^1}{dp}(1-q^*)\frac{\partial p}{\partial q}\Big|_{q^*} - \Delta V^1\right] - \left[\frac{d\Delta V^2}{dp}q^*\frac{\partial p}{\partial q}\Big|_{q^*} + \Delta V^2\right].$$

This is rewritten as

$$\frac{\partial S}{\partial q}\Big|_{q^*} = \left[q^*\Delta V^2(p^*)\left[\frac{1}{\Delta V^1}\frac{d\Delta V^1}{dp} - \frac{1}{\Delta V^2}\frac{d\Delta V^2}{dp}\right]\frac{\partial p}{\partial q}\Big|_{q^*} - \left(\Delta V^1(p^*) + \Delta V^2(p^*)\right)\right] < 0.$$

Now, the condition $\phi'_q(\beta, q) < 4$ implies

$$\left[\frac{1}{\Delta V^1}\frac{d\Delta V^1}{dp} - \frac{1}{\Delta V^2}\frac{d\Delta V^2}{dp}\right]\cdot\frac{\partial p}{\partial q} < 4.$$

This implies

$$\begin{aligned}\frac{\partial S}{\partial q}\Big|_{q^*} &< \left[4q^*\Delta V^2(p^*) - \left(\Delta V^1(p^*) + \Delta V^2(p^*)\right)\right]\\ &= \Delta V^2(p^*)\left[4q^* - 1 - \frac{\Delta V^1(p^*)}{\Delta V^2(p^*)}\right]\\ &= \Delta V^2(p^*)\left[4q^* - 1 - \frac{q^*}{1-q^*}\right]\\ &= \frac{\Delta V^2(p^*)}{1-q^*}[4q^*(1-q^*) - 1] < 0.\end{aligned}$$

Hence, for all point q^* satisfying $S(\beta, q^*) = 0$, one has $(\partial S/\partial q)|_{q^*} < 0$. Given that $S(\beta, 0) = \Delta V^1(p(\beta, 0) > 0$ and $S(\beta, 1) = -\Delta V^2(p(\beta, 1) < 0$, this implies the uniqueness of $q^*(\beta)$ for every value of β, such that $S(\beta, q) > 0$ when $q < q^*(\beta)$ and $S(\beta, q) < 0$ when $q > q^*(\beta)$.

Note that the conditions for this result are in particular satisfied when $\Delta V^1(p)/\Delta V^2(p)$ is a decreasing (increasing) function of the policy p and the equilibrium policy $p(\beta, q)$ is increasing (decreasing) in q. Namely, this happens when an increase in the frequency of a cultural trait induces a change of equilibrium policy that tends to reduce the relative marginal incentives (i.e., paternalistic motive) of family transmission of that trait in the population. QED.

B4.4. Linearized Joint Dynamics under Bisin and Verdier (2001)

We consider cultural transmission under the Bisin and Verdier (2001) model with assumption 3. Assumption 3 implies that $q(\beta) = \hat{q}(p)$ with $p = p(\beta, q)$, with some notational abuse, where $\hat{q}(p) \in [0, 1]$ is the unique solution of the following equation:

$$\frac{\Delta V^1(p)}{\Delta V^2(p)} = \frac{q}{1-q}. \tag{24}$$

The separability of preference structures in assumption 3 implies that the policy instrument p affects the optimal private actions, a^i, independent of the economy-level aggregates **a** and q. This in turn implies that cultural intolerances ΔV^i

depend on only the equilibrium policy level p. As usual, we denote the partial derivative of a variable x on another variable y as $\partial x/\partial y = x_y$.

The linearized local dynamics around the interior steady state (β^*, q^*) of (14) can then easily be obtained by

$$\begin{pmatrix} \dot{\beta} \\ \dot{q} \end{pmatrix} = \begin{pmatrix} \lambda\left[\frac{p_\beta^{\mathrm{com}} - p_\beta}{p_\beta}\right]_{(\beta^*,q^*)} & \lambda\left[\frac{p_q^{\mathrm{com}} - p_q}{p_\beta}\right]_{(\beta^*,q^*)} \\ -\mu G q^*(1-q^*)\cdot \hat{q}_p \cdot p_\beta & \mu G q^*(1-q^*)\left[1 - \hat{q}_p \cdot p_q\right] \end{pmatrix} \begin{pmatrix} \beta \\ q \end{pmatrix}, \tag{25}$$

where $G = -(\Delta V^1(p(\beta^*, q^*)) + \Delta V^2(p(\beta^*, q^*)) < 0$.

The local stability of the interior steady state (β^*, q^*) of (14) is obtained under the standard Hessian conditions:

$$\left[\frac{p_\beta - p_\beta^{\mathrm{com}}}{p_\beta}\right]_{(\beta^*,q^*)} > 0,$$

$$1 - \left[p_q \cdot \hat{q}_p\right]_{(\beta^*,q^*)} > 0, \tag{26}$$

$$\left[(1 - p_q \cdot \hat{q}_p)\cdot\left[\frac{p_\beta - p_\beta^{\mathrm{com}}}{p_\beta}\right] + \hat{q}_p \cdot (p_q - p_q^{\mathrm{com}})\right]_{(\beta^*,q^*)} > 0.$$

The following lemma characterizes the conditions for institutional and cultural dynamics to be complementary or substitute at a locally stable interior steady state.

LEMMA B5. With cultural evolution according to Bisin and Verdier (2001) and under assumption 3, institutional and cultural dynamics are complementary at a locally stable interior steady state (β^*, q^*) if

$$\frac{dP(\beta^*, q^*)}{dq} \text{ has the same sign as } \left[\frac{d\left(\frac{\Delta V^1(p)}{\Delta V^2(p)}\right)}{dp}\right]_{p(\beta^*,q^*)}; \tag{27}$$

they are instead substitute if the signs are opposite.

Proof. Institutional and cultural dynamics are complementary at (β^*, q^*) when

$$\frac{d\beta(q)}{dq} \text{ and } \frac{dq(\beta)}{d\beta} \text{ have the same sign.} \tag{28}$$

Differentiating, we have

$$\frac{d\beta(q)}{dq} = -\frac{(p_q - p_q^{\mathrm{com}})}{p_\beta - p_\beta^{\mathrm{com}}}, \quad \frac{dq(\beta)}{d\beta} = \frac{\hat{q}_p p_\beta}{1 - p_q \cdot \hat{q}_p}.$$

Thus, condition (28) is equivalent to

$$\frac{d\beta(q)}{dq} \cdot \frac{dq(\beta)}{d\beta} \geq 0$$

or

$$-\frac{\left(p_q - p_q^{\mathrm{com}}\right)}{p_\beta - p_\beta^{\mathrm{com}}} \cdot \frac{\hat{q}_p p_\beta}{1 - p_q \cdot \hat{q}_p} \geq 0.$$

Given the Hessian conditions for local stability, (26), at an interior locally stable steady state (β^*, q^*), this condition is equivalent to

$$\left[\left(p_q^{\mathrm{com}} - p_q\right) \cdot \hat{q}_p\right]_{(\beta^*, q^*)} \geq 0\,.$$

Recalling that the cultural manifold $q(\beta)$ is obtained from

$$\frac{\Delta V^1(p)}{\Delta V^2(p)} = \frac{q}{1-q} \text{ and } p = p(\beta, q)$$

and that $P(\beta, q) \coloneqq p^{\mathrm{com}}(\beta, q) - p(\beta, q)$, differentiating, we have

$$\left[\left(p_q^{\mathrm{com}} - p_q\right) \cdot \hat{q}_p\right]_{(\beta^*, q^*)} = \left[P_q(\beta, q) \cdot \left[\frac{d(\Delta V^1(p)/\Delta V^2(p))}{dp}\right]_{p(\beta, q)} (1-q)^2\right]_{(\beta^*, q^*)}.$$

Therefore, institutional and cultural dynamics are complementary at a locally stable interior steady state (β^*, q^*) when P_q and $(d(\Delta V^1(p)/\Delta V^2(p)))/dp$ have the same sign at (β^*, q^*). Obviously, they are dynamic substitute otherwise. QED

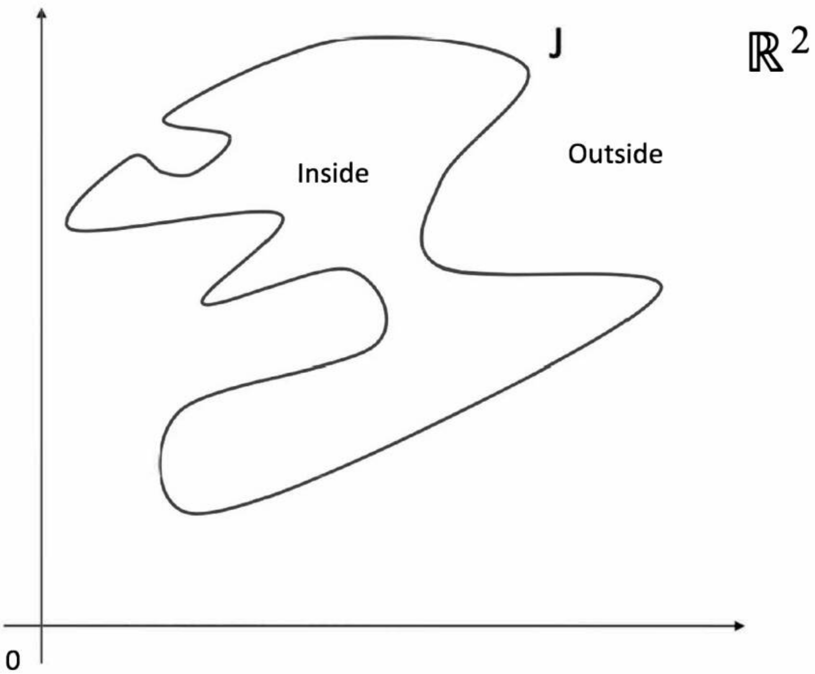


FIG. B1.—Jordan curve J in (nonnegative) plane.

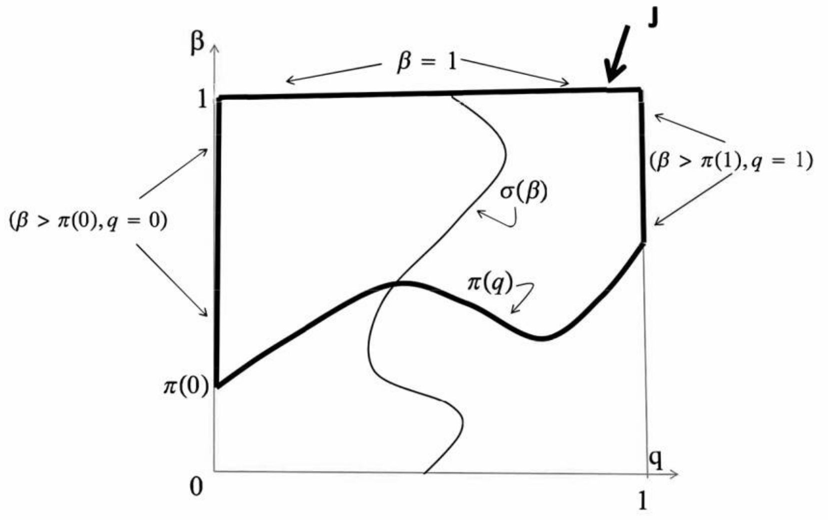


Fig. B2.—Jordan curve J in $[0, 1]^2$, as constructed in proof.

Appendix C

Assumptions on Fundamentals

In this appendix, we translate assumptions 1 and 2 into restrictions on fundamentals.

C1. Sufficient Conditions for the Existence and Monotonicity of the Societal Equilibrium $p(\beta, q)$

Without loss of generality, restrict $p \in [0, 1]$. The indirect utility function can be written as

$$u(\mathbf{a}, p, q) = u(a, p; A, q),$$
$$u^i(\mathbf{a}, p, q) = u^i(a^i, p; A, q) \text{ for } i = 1, 2,$$

where the individual private action $a, a^1, a^2 \in [0, 1]$ and A is an aggregate population-level index $A = A(a, a^1, a^2, p, q)$.

Assume that $u(\cdot)$ ($u^i(\cdot)$ for $i = 1, 2$) is twice differentiable in $(a, p; A, q)$ ($(a^i, p; A, q)$) and strictly concave in a (a^i), that is, $u_{11} < 0$ ($u^i_{11} < 0$). Assume also that the aggregator function $A(\cdot)$ is differentiable in (a, a^1, a^2, p, q) and such that the image of $[0, 1]^5$ by $A(\cdot)$ is an interval $[A_{\min}; A_{\max}]$. Finally, assume the following boundary conditions:

$$u_1(0, p; A, q) \geq 0,\ u_1(1, p; A, q) \leq 0 \text{ for all } (p, A, q) \in [0, 1] \times [A_{\min}; A_{\max}] \times [0, 1]$$
$$u_1^i(0, p; A, q) \geq 0,\ u_1^i(1, p; A, q) \leq 0 \text{ for all } (p, A, q) \in [0, 1] \times [A_{\min}; A_{\max}] \times [0, 1],\ i = 1, 2.$$

These conditions and the fact that $u(\cdot)$ $(u^i(\cdot))$ is a strictly concave function in a (a^i) ensure that the optimal individual behavior for a given value of p and A is characterized by a continuous function (p, A, q) $(a^i(p, A, q)) \in [0, 1]$ obtained from the first-order condition:

$$u_1(a, p; A, q) = 0,$$
$$u_1^i(a^i, p; A, q) = 0 \text{ for } i = 1, 2.$$

For given values of $p \in P$ and $q \in [0, 1]$, a Nash equilibrium in private actions a^N, a^{1N}, a^{2N} and aggregate index A^N (p, q) is characterized by the solution of the following system:

$$a^N = a(p, A, q),\ a^{iN} = a^i(p, A^N, q) \text{ for } i \in (1, 2) \text{ and } A^N = A(a, a^{1N}, a^{2N}, p, q),$$

which in turn translates into the following condition for A^N:

$$A^N = A(a(p, A^N, q), a^1(p, A^N, q), a^2(p, A^N, q), p, q). \tag{29}$$

The following sufficient conditions ensure the existence of a unique Nash equilibrium in private actions a^N (p, q), a^{1N} (p, q), a^{2N} (p, q), A^N (p, q):

$$1 - A'_a \frac{u_{13}}{-u_{11}} - \sum_{i=1,2} A'_{a^i} \frac{u_{13}^i}{-u_{11}^i} > 0 \text{ for all } (a, a^1, a^{2,}, A, p, q),$$
$$A(a(p, A_{\min}, q), a^1(p, A_{\min}, q), a^2(p, A_{\min}, q), q) > A_{\min} \text{ for all } (p, q) \in [0, 1]^2,$$
$$A(a(p, A_{\max}, q), a^1(p, A_{\max}, q), a^2(p, A_{\max}, q), q) < A_{\max} \text{ for all } (p, q) \in [0, 1]^2,$$

with $A'_a = \partial A/\partial a$ and $A'_{a^j} = \partial A/\partial a^j$. The first condition ensures that the function $\Gamma(x, p, q) = x - A(a(p, x, q), a^1(p, x, q), a^2(p, x, q), p, q)$ is increasing for all $(p, q) \in [0, 1]^2$. The second and third conditions ensure that $\Gamma(A_{\min}, p, q) < 0 < \Gamma(A_{\max}, p, q)$. Together these conditions ensure the existence of a unique value A^N (p, q) satisfying (29) and thus correspondingly a unique Nash equilibrium profile a^N (p, q), a^{1N} (p, q), a^{2N} (p, q).

Moreover, differentiating, we have

$$\frac{dA^N}{dp} = \frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2} A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2} A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]},$$

$$\frac{da^N}{dp} = \frac{u_{12}}{-u_{11}} + \frac{u_{13}}{-u_{11}} \frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2} A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2} A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]},$$

$$\frac{da^{iN}}{dp} = \frac{u^i_{12}}{-u^i_{11}} + \frac{u^i_{13}}{-u^i_{11}} \frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2} A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2} A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]}.$$

The condition for an interior societal equilibrium $p(\beta, q)$ is obtained from the first-order conditions of the policy maker,

$$(1-\beta)\cdot u_2(a, p, A, q) + \beta\left[q \cdot u^1_2\left(a^1, p, A, q\right) + (1-q)\cdot u^2_2(a^2, p, A, q)\right] = 0\,.$$

After substitution of the Nash equilibrium private actions $a^N\,(p, q)$, $a^{1N}\,(p, q)$, $a^{2N}\,(p, q)$, $A^N\,(p, q)$, this condition can be written as

$$\Psi(p, q, \beta) = 0, \tag{30}$$

with

$$\begin{aligned}\Psi(p, q, \beta) = {} & (1-\beta)u_2(a^N(p, q), p, A^N(p, q), q) \\ & + \beta\cdot\begin{bmatrix} q\cdot u^1_2\left(a^{1N}(p, q), p, A^N(p, q), q\right) \\ +(1-q)\cdot u^2_2(a^{2N}(p, q), p, A^N(p, q), q)\end{bmatrix}.\end{aligned}$$

Moreover, a corner societal equilibrium $p(\beta, q) = 0$ ($p(\beta, q) = 1$) obtains when $\Psi(0, q, \beta) \leq 0$ ($\Psi(1, q, \beta) \geq 1$).

A sufficient condition for the existence of a unique societal equilibrium $p(\beta, q)$ consists in the function $\Psi\,(p, q, \beta)$ being decreasing in p for all $q \in [0, 1]$. Given the smoothness assumptions on the functions $u(\cdot)$, $u^i(\cdot)$, and $A(\cdot)$, this is satisfied when the following condition holds:

$$\begin{aligned} & u_{12}\frac{da^N}{dp} + u_{22} + u_{23}\frac{dA^N}{dp} < 0, \\ & u^i_{12}\frac{da^{iN}}{dp} + u^i_{22} + u^i_{23}\frac{dA^N}{dp} < 0 \text{ for all } i \in (1, 2)\,. \end{aligned}$$

In turn, in terms of the fundamentals, this conditions becomes

$$\frac{u_{12}}{-u_{22}}\left[\frac{u_{12}}{-u_{11}}+\frac{u_{13}}{-u_{11}}\frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2}A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2}A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]}\right]$$
$$+\frac{u_{23}}{-u_{22}}\frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2}A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2}A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]} < 1,$$

and for $i = 1, 2$,

$$\frac{u^i_{12}}{-u^i_{22}}\left[\frac{u^i_{12}}{-u^i_{11}}+\frac{u^i_{13}}{-u^i_{11}}\frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2}A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2}A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]}\right]$$
$$+\frac{u^i_{23}}{-u^i_{22}}\frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2}A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2}A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]} < 1,$$

or

$$\frac{(u_{12})^2}{u_{22}u_{11}}+\left(\frac{u_{12}}{u_{22}}\frac{u_{13}}{u_{11}}+\frac{u_{23}}{(-u_{22})}\right)\frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2}A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2}A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]} < 1,$$
$$\frac{(u^i_{12})^2}{u^i_{22}u^i_{11}}+\left(\frac{u^i_{12}}{u^i_{22}}\frac{u^i_{13}}{u^i_{11}}+\frac{u^i_{23}}{(-u^i_{22})}\right)\frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2}A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2}A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]} < 1$$
$$\text{for } i \in (1, 2),$$

(recalling $A'_p = \partial A/\partial p$, $A'_a = \partial A/\partial a$, and $A'_{a^j} = \partial A/\partial a^j$). These conditions are more likely to be satisfied when $|u_{11}|$, $|u^i_{11}|$, and $|u^i_{22}|$ are large enough.

This condition simplifies if the preferences structure is characterized by some degree of separability:

$$u(a, p; A, q) = v(a, p, \theta) + H(p, A),$$
$$u^i(a, p; A, q) = v(a, p, \theta_i) + H(p, A),$$

and $\theta_i > \theta$ for $i = 1, 2$. Such preferences lead to $a^N = a(p, \theta)$, $a^{1N} = a(p, \theta_1)$, $a^{2N} = a(p, \theta_2)$, and $A^N = A(a(p, \theta), a(p, \theta_1), a(p, \theta_2), p, q)$.

A sufficient condition for the existence of a unique societal equilibrium, given that $u_{13} = 0$ and $u^j_{13} = 0$, is then

$$\frac{(v_{12})^2}{\left[v_{22} + H_{pp}\right]v_{11}}+\left(\frac{H_{pA}}{-(v_{22} + H_{pp})}\right)\left[A'_p + A'_a\frac{v_{12}}{-v_{11}} + \sum_{j=1,2}A'_{a^j}\frac{v^j_{12}}{-v^j_{11}}\right] < 1,$$
$$\frac{(v^i_{12})^2}{\left[v^i_{22} + H_{pp}\right]v^i_{11}}+\left(\frac{H_{pA}}{-(v^i_{22} + H_{pp})}\right)\left[A'_p + A'_a\frac{v_{12}}{-v_{11}} + \sum_{j=1,2}A'_{a^j}\frac{v^j_{12}}{-v^j_{11}}\right] < 1 \text{ for } i \in (1, 2),$$

where $v_{kl} = v''_{kl}(a, p, \theta)$ and $v^i_{kl} = v''_{kl}(a, p, \theta_i)$. But

$$\begin{aligned} u_2^i - u_2 &= u_2(a^{1N}, p, A^N, q, \theta_i) - u_2(a^N, p, A^N, q, \theta) \\ &= v_2(a(p, \theta_i), p, \theta_i) - v_2(a(p, \theta), p, \theta)\,. \end{aligned}$$

For instance, consider as an example the following preference structure:

$$\begin{aligned} u(a, p, A, q) &= (1-p)a + \theta W(1-a) + H(p, A), \\ u^i(a, p, A, q) &= (1-p)a + \theta^i W(1-a) + H(p, A), \end{aligned}$$

with $W(\cdot)$ a strictly increasing and concave function, $A = \lambda a + (1-\lambda)[qa^1 + (1-q)a]^2$, and $H(p, A)$ concave in p. Then,

$$\begin{aligned} v_a &= (1-p) - \theta W'(1-a),\ v_p = -a \\ v_{ap} &= -1, v_{a\theta} = -W'(1-a), \\ -v_{aa} &= -\theta W''(1-a), \\ v_{p\theta} &= 0,\ v_{pp} = 0\,. \end{aligned}$$

The sufficient condition for a well-defined societal equilibrium $p(\beta, q)$ can be written as

$$\begin{aligned} &\frac{1}{H_{pp}\theta W''} + \left(\frac{H_{pA}}{-(H_{pp})}\right)\left[\lambda\frac{1}{\theta W''} + (1-\lambda)\left[q\frac{1}{\theta_1 W''(1)} + (1-q)\frac{1}{\theta_2 W''(2)}\right]\right] < 1, \\ &\frac{1}{H_{pp}\theta_i W''} + \left(\frac{H_{pA}}{-(H_{pp})}\right)\left[\lambda\frac{1}{\theta W''} + (1-\lambda)\left[q\frac{1}{\theta_1 W''(1)} + (1-q)\frac{1}{\theta_2 W''(2)}\right]\right] \\ &\qquad < 1 \text{ for } i = 1, 2, \end{aligned}$$

where $W'' = W''(1-a^N)$, $W''(i) = W''(1-a^{Ni})$ for $i = 1, 2$.

When $H_{pA} > 0$, given that $-(H_{pp}) > 0$ and $[\lambda(1/\theta W'') + (1-\lambda)[q(1/\theta_1 W''(1)) + (1-q)(1/\theta_2 W''(2))]] < 0$, this condition is satisfied when $(1/H_{pp}\theta W'') < 1$ and $(1/H_{pp}\theta_i W''(i)) < 1$, which in turn holds when $1 < H_{pp} \min\{W''\theta, \theta_1 W''(1), \theta_2 W''(2)\}$. This is satisfied with enough concavity of W and H, respectively, in a and p.

When $H_{pA} < 0$, this sufficient condition can be rewritten as

$$\frac{1}{\theta} - H_{pA}\left[\lambda\frac{1}{\theta} + (1-\lambda)\left[q\frac{1}{\theta_1} + (1-q)\frac{1}{\theta_2}\right]\right] < H_{pp} W'',$$

which again will be satisfied when H_{pA} is bounded from below on the relevant domain, $[0, 1] \times [A_{\min}, A_{\max}]$ (i.e., $H_{pA} > -K$, with $K > 0$) and $H_{pp} W'' > (1 + K)/\theta$. This is also satisfied with enough concavity of W and H, respectively, in a and p.

Monotonicity of the societal equilibrium $p(\beta, q)$.—Once we have the existence of a unique smooth societal equilibrium $p(\beta, q)$, the monotonicity in β is easy to show. Indeed, differentiating (30), we obtain

$$\frac{\partial p}{\partial \beta} = \frac{\Psi_\beta}{-\Psi_p} = \frac{[q \cdot u_2^1 + (1-q) \cdot u_2^2] - u_2}{-\Psi_p}\,.$$

As $p(\beta, q)$ is the unique solution of (30), one should have $\Psi_p < 0$ at the point $p(\beta, q)$. Thus, the sign of $p(\beta, q)$ is the same as the sign of the partial derivative Ψ_β

evaluated at the point $p(\beta, q)$. Since the function $\Psi(p, q, \beta)$ is linear in β, this partial derivative $\Psi_\beta = [q \cdot u_2^1 + (1-q) \cdot u_2^2] - u_2$ is a function of β only via $p = p(\beta, q)$. Let us call this partial derivative $\Psi_\beta = f(p(\beta, q))$.

Now suppose that $p(\beta, q)$ is nonmonotonic, and suppose, for instance, that there is a local strict maximum of $p(\beta, q)$ at some point β_0, with $p(\beta_0, q) = A$. We have $\partial p/\partial\beta = 0$ at β_0, and therefore $f(A) = 0$. Also, we should have $\partial^2 p/\partial\beta^2 < 0$ at β_0. Now it is easy to see that at β_0, $\partial^2 p/\partial\beta^2 = f'(A)f(A)/\Psi_p^2 = 0$, therefore leading to a contradiction.

One can of course do the same reasoning if we take a local strict minimum for $p(\beta, q)$.

It follows that once we have sufficient conditions for a unique smooth societal equilibrium $p(\beta, q)$, it is necessarily monotonic in β.

C2. Sufficient Conditions for the Existence of the Societal Commitment Equilibrium $p^{\mathrm{com}}(\beta, q)$

The societal commitment equilibrium given institutions β and cultural distribution q is obtained from the following maximization problem:

$$\max\ (1-\beta)u(a^N, p; A^N, q) + \beta\left[q \cdot u^1\left(a^{1N}, p; A^N, q\right) + (1-q) \cdot u^2\left(a^{2^N}, p; A^N, q\right)\right]$$
$$\text{such that } a^N = a^N(p, q),\ a^{iN} = a^{iN}(p, q) \text{ for } i \in (1, 2),$$
$$\text{and } A^N = A^N(p, q).$$

Let

$$\Omega(p, \beta, q) = (1-\beta)u(a^N(p, q), p; A^N(p, q), q) + \beta\begin{bmatrix} q \cdot u^1\left(a^{1N}(p, q), p; A^N(p, q), q\right) \\ +(1-q) \cdot u^2(a^{2N}(p, q), p; A^N(p, q), q) \end{bmatrix}.$$

The first-order condition for an interior societal commitment equilibrium $p^{\mathrm{com}}(\beta, q)$ can be written as

$$\begin{aligned}\Omega_p(p, \beta, q) &= (1-\beta)\left[u_2(a^N, p; A^N, q) + u_3(a^N, p, A^N, q)\frac{dA^N}{dp}\right] \\ &\quad + \beta\begin{bmatrix} q\left(u_2^1(a^{1N}, p, A^N, q) + u_3^1(a^{1N}, p, A^N, q)\dfrac{dA^N}{dp}\right) \\ +(1-q)\left(u_2^2(a^{1N}, p, A^N, q) + u_3^2(a^{1N}, p, A^N, q)\dfrac{dA^N}{dp}\right)\end{bmatrix} \\ &= 0.\end{aligned}$$

A sufficient (strong) condition is then

$$\Omega_{pp}(p, \beta, q) < 0 \text{ for all } p \in [0, 1].$$

Differentiating, we have

$$\begin{aligned}
\Omega_{pp}(p,\beta,q) \;=\; &(1-\beta)\left[u_{12}\frac{da^N}{dp}+u_{22}+u_{23}\frac{dA^N}{dp}\right]\\
&+\beta\left[\begin{array}{c} q\left(u^1_{12}\dfrac{da^{1N}}{dp}+u^1_{22}+u^1_{23}\dfrac{dA^N}{dp}\right)\\ +(1-q)\left(u^2_{12}\dfrac{da^{2N}}{dp}+u^2_{22}+u^2_{23}\dfrac{dA^N}{dp}\right)\end{array}\right]\\
&+(1-\beta)\left[u_{13}\frac{da^{iN}}{dp}+u_{23}+u_{33}\right]\frac{dA^N}{dp}\\
&+\beta\left[\begin{array}{c} q\left[u^1_{13}\dfrac{da^{1N}}{dp}+u^1_{23}+u^1_{33}\right]\dfrac{dA^N}{dp}\\ +(1-q)\left[u^2_{13}\dfrac{da^{2N}}{dp}+u^2_{23}+u^2_{33}\right]\dfrac{dA^N}{dp}\end{array}\right]\\
&+(1-\beta)u_3\frac{d^2A^N}{dp^2}+\beta\left[\left(qu^1_3+(1-q)u^2_3\right)\frac{d^2A^N}{dp^2}\right].
\end{aligned}$$

Thus, a sufficient condition for $\Omega_{pp}(p,\beta,q)<0$ is that

$$u_{12}\frac{da^N}{dp}+u_{22}+u_{23}\frac{dA^N}{dp}+\left[u_{13}\frac{da^N}{dp}+u_{23}+u_{33}\right]\frac{dA^N}{dp}+u_3\frac{d^2A^N}{dp^2}<0,$$

and for $i=1,2$,

$$u^i_{12}\frac{da^{iN}}{dp}+u^i_{22}+u^i_{23}\frac{dA^N}{dp}+\left[u^i_{13}\frac{da^{iN}}{dp}+u^i_{23}+u^i_{33}\right]\frac{dA^N}{dp}+u^i_3\frac{d^2A^N}{dp^2}<0.$$

Recall

$$\begin{aligned}
\frac{da^N}{dp}&=\frac{u_{12}}{-u_{11}}+\frac{u_{13}}{-u_{11}}\frac{A'_p+A'_a(u_{12}/-u_{11})+\sum_{j=1,2}A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1-A'_a(u_{13}/-u_{11})-\sum_{j=1,2}A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]},\\
\frac{da^{iN}}{dp}&=\frac{u^i_{12}}{-u^i_{11}}+\frac{u^i_{13}}{-u^i_{11}}\frac{A'_p+A'_a(u_{12}/-u_{11})+\sum_{j=1,2}A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1-A'_a(u_{13}/-u_{11})-\sum_{j=1,2}A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]},
\end{aligned}$$

and

$$\frac{dA^N}{dp}=\frac{A'_p+A'_a(u_{12}/-u_{11})+\sum_{j=1,2}A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1-A'_a(u_{13}/-u_{11})-\sum_{j=1,2}A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]}.$$

Tedious manipulations show then that a sufficient condition for $\Omega_{pp}(p,\beta,q)<0$ is that

$$
\begin{aligned}
D = {} & \frac{(u_{12})^2}{-u_{11}} + u_{22} \\
& + \left(2\left(\frac{u_{13}u_{12}}{-u_{11}} + u_{23}\right) + u_{33}\right) \frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2} A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2} A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]} \\
& + \frac{(u_{13})^2}{-u_{11}} \left[\frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2} A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2} A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]}\right]^2 + u_3 \frac{d^2 A^N}{dp^2} < 0,
\end{aligned}
$$

and that for $i = 1, 2$,

$$
\begin{aligned}
D^i = {} & \frac{(u^i_{12})^2}{-u^i_{11}} + u^i_{22} \\
& + \left(2\left(\frac{u^i_{13}u^i_{12}}{-u^i_{11}} + u^i_{23}\right) + u^i_{33}\right) \frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2} A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2} A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]} \\
& + \frac{(u^i_{13})^2}{-u^i_{11}} \left[\frac{A'_p + A'_a(u_{12}/-u_{11}) + \sum_{j=1,2} A'_{a^j}\left(u^j_{12}/-u^j_{11}\right)}{\left[1 - A'_a(u_{13}/-u_{11}) - \sum_{j=1,2} A'_{a^j}\left(u^j_{13}/-u^j_{11}\right)\right]}\right]^2 + u^i_3 \frac{d^2 A^N}{dp^2} < 0.
\end{aligned}
$$

Because of the term in d^2A^N/dp^2, this involves complicated conditions on the third derivatives of the indirect preference functions. When preferences are separable of the form

$$
\begin{aligned}
u(a, p; A, q) &= v(a, p, \theta) + H(p, A), \\
u^i(a, p; A, q) &= v(a, p, \theta^i) + H(p, A) \text{ for } i = 1, 2,
\end{aligned}
$$

the expression D and D^i simplifies somewhat:

$$
\begin{aligned}
D &= \frac{(v_{ap})^2}{-v_{aa}} + v_{pp} + (2H_{pA} + H_{AA})\left(A_p + A'_a \frac{v_{ap}}{-v_{pp}} + \sum_{j=1,2} A'_{a^j} \frac{v^j_{ap}}{-v^j_{pp}}\right) + H_A \frac{d^2 A^N}{dp^2}, \\
D^i &= \frac{(v^i_{ap})^2}{-v^i_{aa}} + v^i_{pp} + (2H_{pA} + H_{AA})\left(A_p + A'_a \frac{v_{ap}}{-v_{pp}} + \sum_{j=1,2} A'_{a^j} \frac{v^j_{ap}}{-v^j_{pp}}\right) + H_A \frac{d^2 A^N}{dp^2}.
\end{aligned}
$$

Therefore, Ω (p, β, q) is strictly concave in p when $v(a, p, \theta)$ and v (a, p, θ^i) are sufficiently concave in (a, p).

References

Arroyo Abad, L., and N. Maurer. 2021. “History Never Really Says Goodbye: A Critical Review of the Persistence Literature.” *J. Hist. Polit. Econ.* 1 (1): 31–68.

Acemoglu, D. 2003. "Why Not a Political Coase Theorem? Social Conflict, Commitment, and Politics." *J. Comparative Econ.* 31:620–52.

Acemoglu, D., G. Egorov, and K. Sonin. 2015. "Political Economy in a Changing World." *J.P.E.* 123 (5): 1038–86.

———. 2021. "Institutional Change and Institutional Persistence." In *Handbook of Historical Economics*, edited by A. Bisin and G. Federico, 365–89. Amsterdam: Elsevier.

Acemoglu, D., and M. Jackson. 2015. "History, Expectations, and Leadership in the Evolution of Social Norms." *Rev. Econ. Studies* 82 (2): 423–56.

Acemoglu, D., S. Johnson, and J. A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *A.E.R.* 91:1369–401.

———. 2006. "Institutions as the Fundamental Cause of Long-Run Growth." In *Handbook of Economic Growth*, edited by P. Aghion and S. Durlauf. Amsterdam: Elsevier.

Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared. 2008. "Income and Democracy." *A.E.R.* 98 (3): 808–42.

Acemoglu, D., and J. A. Robinson. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Q.J.E.* 115 (4): 1167–99.

———. 2001. "A Theory of Political Transitions." *A.E.R.* 91 (4): 938–63.

———. 2006. *Economic Origins of Dictatorship and Democracy*. Cambridge: Cambridge Univ. Press.

———. 2010. *Why Nations Fail*: New York: Crown.

———. 2021. "Culture, Institutions and Social Equilibria: A Framework." Working Paper no. 28832, NBER, Cambridge, MA.

Aghion, P., I Algan, P. Cahuc, and A. Schleifer. 2010. "Regulation and Distrust." *Q.J.E.* 125 (3): 1015–49.

Alesina, A., and P. Giuliano. 2015. "Culture and Institutions." *J. Econ. Literature* 53 (4): 898–944.

Alexander, J. C. 2003. *The Meanings of Social Life: A Cultural Sociology*. New York: Oxford Univ. Press.

Alger, I., and J. W. Weibull. 2013. "Homo Moralis—Preference Evolution under Incomplete Information and Assortative Matching." *Econometrica* 8:2269–302.

Almagro, M., and D. Andres-Cerezo. 2020. "The Construction of National Identities." *Theoretical Econ.* 15 (2): 763–810.

Aron, R. 1950a. "Social Structure and the Ruling Class: Part 1." *British J. Sociology* 1 (1): 1–16.

———. 1950b. "Social Structure and the Ruling Class: Part 2." *British J. Sociology* 1 (2): 126–43.

Arroyo Abad, L., and J. Luiten van Zanden. 2016. "Growth under Extractive Institutions? Latin American Per Capita GDP in Colonial Times." *J. Econ. Hist.* 76 (4): 1182–215.

Belloc, M., and S. Bowles. 2013. "The Persistence of Inferior Cultural-Institutional Conventions." *A.E.R. Papers and Proc.* 103 (3): 1–7.

———. 2017. "Persistence and Change in Culture and Institutions under Autarchy, Trade, and Factor Mobility." *American Econ. J. Microeconomics* 9 (4): 245–76.

Benabou, R., D. Ticchi, and A. Vindigni. 2015. "Religion and Innovation." *A.E.R.* 105 (5): 346–51.

Benhabib, J., A. Corvalan, and M. M. Spiegel. 2013. "Income and Democracy: Evidence from Nonlinear Estimations." *Econ. Letters* 118:489–92.

Berman, S. 1997. "Civil Society and the Collapse of the Weimar Republic." *World Politics* 49:401–29.

Besley, T. 2017. "Aspirations and the Political Economy of Inequality." *Oxford Econ. Papers* 69 (1): 1–35.

———. 2020. "State Capacity, Reciprocity, and the Social Contract." *Econometrica* 88 (4): 1307–35.

Besley, T., and M. Ghatak. 2017. "The Evolution of Motivation." Working paper, London School Econ. and Polit. Sci.

Besley, T., and T. Persson. 2019. "Democratic Values and Institutions." *A.E.R. Insights* 1 (1): 59–76.

———. 2020. "Organizational Dynamics: Culture, Design, and Performance." Working paper, London School Econ. and Polit. Sci.

Bidner, C., and P. Francois. 2011. "Cultivating Trust: Norms, Institutions and the Implications of Scale." *Econ. J.* 121 (555): 1097–129.

Bisin, A., and G. Federico, eds. 2021. *Handbook of Historical Economics*. Amsterdam: Elsevier.

Bisin, A., and A. Moro. 2021. "LATE for History." In *Handbook of Historical Economics*, edited by A. Bisin and G. Federico, 269–96. Amsterdam: Elsevier.

Bisin, A., J. Rubin, A. Seror, and T. Verdier. 2021. "Culture, Institutions and the Long Divergence." Working Paper no. 28488, NBER, Cambridge, MA.

Bisin, A., and T. Verdier. 1998. "On the Cultural Transmission of Preferences for Social Status." *J. Public Econ.* 70:75–97.

———. 2000a. "Beyond the Melting Pot: Cultural Transmission, Marriage and the Evolution of Ethnic and Religious Traits." *Q.J.E.* 115:955–88.

———. 2000b. "Models of Cultural Transmission, Voting and Political Ideology." *European J. Polit. Econ.* 16:5–29.

———. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *J. Econ. Theory* 97:298–319.

———. 2011. "The Economics of Cultural Transmission and Socialization." In *Handbook of Social Economics*, edited by J. Benhabib, A. Bisin, and M. Jackson, 339–416. Amsterdam: Elsevier.

———. 2021. "Phase Diagrams in Historical Economics: Culture and Institutions." In *Handbook of Historical Economics*, edited by A. Bisin and G. Federico, 491–522. Amsterdam: Elsevier.

Bjornerstedt, J., and J. Weibull. 1996. "Nash Equilibrium and Evolution by Imitation." In *The Rational Foundations of Economic Behaviour*, edited by K. J. Arrow, E. Colombatto, M. Perlman, and C. Schmidt, 155–81. New York: St. Martin's.

Bogart, D., and L. Chaudhary. 2019. "Extractive Institutions? Investor Returns to Indian Railway Companies in the Age of High Imperialism." *J. Inst. Econ.* 15 (5): 751–74.

Boranbay, S., and C. Guerriero. 2019. "Endogenous (In)Formal Institutions." *J. Comparative Econ.* 47 (4): 921–45.

Bowles, S., J.-K. Choi, S.-H. Hwang, and S. Naidu. 2021. "How Institutions and Cultures Change: An Evolutionary Perspective." In *Handbook of Historical Economics*, edited by A. Bisin and G. Federico, 391–433. Amsterdam: Elsevier.

Boyd, R., and P. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: Univ. Chicago Press.

Cantoni, D., and N. Yuchtman. 2021. "Historical Natural Experiments: Bridging Economics and Economic History." In *Handbook of Historical Economics*, edited by A. Bisin and G. Federico, 213–41. Amsterdam: Elsevier.

Carvalho, J. P. 2016. "Identity-Based Organizations." *A.E.R.* 106 (5): 410–14.

Carvalho, J. P., and C. Dippel. 2016. "Elite Identity and Political Accountability: A Tale of Ten Islands." Working Paper no. 22777, NBER, Cambridge, MA.

Carvalho, J. P., and M. Sacks. 2021. "The Economics of Religious Communities." *J. Public Econ.* 201:104481.

Casey, G., and M. Klemp. 2021. "Historical Instruments and Contemporary Endogenous Regressors." *J. Development Econ.* 149:102586.

Cavalli-Sforza, L. L., and M. Feldman. 1973. "Cultural versus Biological Inheritance: Phenotypic Transmission from Parent to Children." *American J. Human Genetics* 25:618–37.

———. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*: Princeton, NJ: Princeton Univ. Press.

Chen, Y., H. Wang, and S. Yan. 2022. "The Long-Term Effects of Protestant Activities in China." *J. Comparative Econ.* 50 (2): 394–414.

Clark, G. 2007. *A Farewell to Alms: A Brief Economic History of the World*. Princeton, NJ: Princeton Univ. Press.

Conley, J., and A. Temimi. 2001. "Endogenous Enfranchisement When Groups' Preferences Conflict." *J.P.E.* 109 (1): 79–102.

Davis, L. S., and C. R. Williamson. 2016. "Culture and the Regulation of Entry." *J. Comparative Econ.* 44 (4): 1055–83.

Dell, M. 2010. "The Persistent Effects of Peru's Mining Mita." *Econometrica* 78 (6): 1863–903.

Dell, M., and B. A. Olken. 2020. "The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java." *Rev. Econ. Studies* 87 (1): 164–203.

Demsetz, H. 1967. "Toward a Theory of Property Rights" *A.E.R. Papers and Proc.* 57:347–59.

Diamond, J. 2005. *Collapse: How Societies Choose to Fail or Succeed*. New York: Penguin.

DiMaggio, P., and H. R. Markus. 2010. "Culture and Social Psychology: Converging Perspectives." *Soc. Psychology Q.* 73:347–52.

Dixit, A., and J. Londregan. 1996. "The Determinants of Success of Special Interests in Redistributive Politics." *J. Politics* 58:1132–55.

Doepke, M., and F. Zilibotti. 2008. "Occupational Choice and the Spirit of Capitalism." *Q.J.E.* 123 (2): 747–93.

Engerman, S., and K. Sokoloff. 2002. "Factor Endowments, Inequality, and Paths of Development among New World Economies." Working Paper no. 9259, NBER, Cambridge, MA.

Fatton, R., Jr. 1995. "Africa in the Age of Democratization: The Civic Limitations of Civil Society." *African Studies Rev.* 38 (2): 67–99.

Finan, F., and L. Schechter. 2012. "Vote-Buying and Reciprocity." *Econometrica* 80 (2): 863–81.

Geertz, C. 1973. *The Interpretation of Cultures.* New York: Basic.

Gorodnichenko, Y., and G. Roland. 2011. "Which Dimensions of Culture Matter for Long-Run Growth?" *A.E.R. Papers and Proc.* 101 (3): 492–98.

———. 2017. "Culture, Institutions, and the Wealth of Nations." *Rev. Econ. and Statis.* 99 (3): 402–16.

———. 2021. "Culture, Institutions and Democratization." *Public Choice* 187 (2): 165–95.

Gradstein, M. 2007. "Inequality, Democracy and the Protection of Property Rights." *Econ. J.* 117 (516): 252–69.

———. 2008. "Institutional Traps and Economic Growth." *Internat. Econ. Rev.* 49 (3): 1043–66.

Greif, A., and G. Tabellini. 2010. "Cultural and Institutional Bifurcation: China and Europe Compared." *A.E.R. Papers and Proc.* 100 (2): 1–10.

———. 2017. "The Clan and the Corporation: Sustaining Cooperation in China and Europe." *J. Comparative Econ.* 45 (1): 1–35.
Grossman, G. M., and E. Helpman. 1994. "Protection for Sale." *A.E.R.* 84 (4): 833–50.
———. 1996. "Electoral Competition and Special Interest Politics." *Rev. Econ. Studies* 63:265–86.
———. 2001. *Special Interest Politics.* Cambridge, MA: MIT Press.
Grossman, S. J., and O. D. Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *J.P.E.* 94 (4): 691–719.
Guimaraes, B., and K. D. Sheedy. 2016. "Guarding the Guardians." *Econ. J.* 127 (606): 2441–77.
Guiso, L., and P. Pinotti. 2012. "Democratization and Civic Capital." Discussion Paper no. 8847, Center Econ. Policy Res., Washington, DC.
Guiso, L., P. Sapienza, and L. Zingales. 2008. "Social Capital and Good Culture." *J. European Econ. Assoc.* 6 (2–3): 295–320.
———. 2016. "Long-Term Persistence." *J. European Econ. Assoc.* 14 (6): 1401–36.
Güth, W. 1995. "An Evolutionary Approach to Explaining Cooperative Behavior by Reciprocal Incentives." *Internat. J. Game Theory* 24:323–44.
Güth, W., and M. Yaari. 1992. "An Evolutionary Approach to Explain Reciprocal Behavior in a Simple Strategic Game." In *Explaining Process and Change—Approaches to Evolutionary Economics*, edited by U. Witt, 23–34. Ann Arbor, MI: Univ. Michigan Press.
Gyimah-Boadi. 2021. "Democratic Backsliding in West Africa: Nature, Causes, Remedies." Working paper, Kofi Annan Found.
Harrison, L. E., and S. P. Huntington, eds. 2001. *Culture Matters: How Values Shape Human Progress.* New York: Basic Books.
Hatcher, A. 2002. *Algebraic Topology.* Cambridge: Cambridge Univ. Press.
Hauk, E., and H. Mueller. 2015. "Cultural Leaders and the Clash of Civilizations." *J. Conflict Resolution* 59 (3): 367–400.
Helbing, D. 1992. "A Mathematical Model for Behavioral Changes by Pair Interactions." In *Economic Evolution and Demographic Change: Formal Models in Social Sciences*, edited by G. Haag, U. Mueller, and K. G. Troitzsch, 330–48. Berlin: Springer.
Huillery, E. 2009. "History Matters: The Long-Term Impact of Colonial Public Investments in French West Africa." *American Econ. J. Applied Econ.* 1 (2): 176–215.
Hofbauer, J. 1995. "Imitation Dynamics for Games." Working paper, Univ. Vienna.
Iyer, L. 2010. "Direct versus Indirect Colonial Rule in India: Long-Term Consequences." *Rev. Econ. and Statis.* 92 (4): 693–713.
Iygun, M., J. Rubin, and A. Seror. 2021. "A Theory of Cultural Revivals." *European Econ. Rev.* 135:103734.
Jack, W., and R. Lagunoff. 2006. "Dynamic Enfranchisement." *J. Public Econ.* 90 (4–5): 551–72.
Kaneko, M., and K. Nakamura. 1979. "The Nash Social Welfare Function." *Econometrica* 47:4233–43.
Kolmogorov, A. 1936. "Sulla Teoria di Volterra sulla Lotta per l'Esistenza." *Giornale dell'Istituto Italiano degli Attuari* 7:78–80.
Lagunoff, R. 2009. "Dynamic Stability and Reform of Political Institutions." *Games and Econ. Behavior* 67 (2): 569–83.
Landes, D. S. 1985. *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor.* New York: W. W. Norton.
Lange, M., J. Mahoney, and M. vom Hau. 2006. "Colonialism and Development: A Comparative Analysis of Spanish and British Colonies." *American J. Sociology* 111 (5): 1412–62.

Levine, D., and S. Modica. 2021. "State Power and Conflict Driven Evolution." In *Handbook of Historical Economics,* edited by A. Bisin and G. Federico, 435–62. Amsterdam: Elsevier.

Lindbeck, A. 1995. "Hazardous Welfare State Dynamics." *A.E.R.* 4:9–15.

Lindbeck, A., and J. Weibull. 1987. "Balanced-Budget Redistribution as the Outcome of Political Competition." *Public Choice* 52:273–97.

Lipset, S. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Polit. Sci. Rev.* 53 (1): 69–105.

Lizzeri, A., and N. Persico. 2004. "Why Did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform.'" *Q.J.E.* 119 (2): 707–65.

Lotka, A. J. 1920. "Analytical Note on Certain Rhythmic Relations in Organic Systems." *Proc. Nat. Acad. Sci. USA* 6 (7): 410–15.

———. 1925. *Elements of Physical Biology.* New York: Williams and Wilkins.

Lowes, S., and E. Monteiro. 2017. "Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State." Working paper, Harvard Univ.

Lowes, S., N. Nunn, J. A. Robinson, and J. Weigel. 2017. "The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom." *Econometrica* 85:1065–91.

Mahoney, J., and K. Thelen, eds. 2010. *Explaining Institutional Change: Ambiguity, Agency, and Power.* Cambridge: Cambridge Univ. Press.

Marx, K. 1867. *Das Kapital, Kritik der Politischen Ökonomie.* Hamburg: von Otto Meisner.

McCloskey, D. N. 2006. *The Bourgeois Virtues: Ethics for an Age of Commerce.* Chicago: Univ. Chicago Press.

———. 2010. *Bourgeois Dignity: Why Economics Can't Explain the Modern World.* Chicago: Univ. Chicago Press.

———. 2017. *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World.* Chicago: Univ. Chicago Press.

Michalopoulos, S., and E. Papaioannou. 2013. "Pre-Colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81 (1): 113–52.

———. 2016. "The Long-Run Effects of the Scramble for Africa." *A.E.R.* 106 (7): 1802–48.

Milnor, J. W. (1965) 1997. "*Topology from the Differentiable Viewpoint.*" Princeton, NJ: Princeton Univ. Press. Original edition: Charlottesville, VA: Univ. Press Virginia.

Mokyr, J. 2016. *A Culture of Growth: The Origins of the Modern Economy.* Princeton, NJ: Princeton Univ. Press.

Mosca, G. 1896. *Elementi di Scienza Politica.* Rome: Fratelli Bocca.

North, D. C., and B. R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *J. Econ. Hist.* 49 (4): 803–32.

Nunn, N. 2007. "Historical Legacies: A Model Linking Africa's Past to Its Current Underdevelopment." *J. Development Econ.* 83:157–75.

———. 2008. "The Long-Term Effects of Africa's Slave Trades." *Q.J.E.* 123:139–76.

Nunn, N., and L. Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *A.E.R.* 101 (7): 3221–52.

Ogilvie, S. 2021. "Thinking Carefully about Inclusiveness: Evidence from European Guilds." *J. Inst. Econ.* 17 (2): 185–200.

Ogilvie, S., and A. W. Carus. 2014. "Institutions and Economic Growth in Historical Perspective." In *Handbook of Economic Growth,* edited by P. Aghion and S. Durlauf, 403–513. Amsterdam: Elsevier.

Paredes-Fuentes, S. 2016. "Origins of Institutions and Inequality in Latin America." Working paper, Univ. Warwick.

Pareto, V. 1901. "Un' applicazione di Teorie Sociologiche." *Rivista Italiana di Sociologia*, 402–56. English translation: 1968. *The Rise and Fall of Elites: An Application of Theoretical Sociology*. London: Bedminster.

———. 1916. *Trattato di sociologia generale*, Florence: G. Barbéra.

Pearsall, E. 1965. "Social Welfare Functions: An Axiomatic Approach." *Zeitschrift für Nationalokonomie/J. Econ.* 25:278–86.

Persson, T., and G. Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge, MA: MIT Press.

———. 2021. "Culture, Institutions and Policy." In *Handbook of Historical Economics*, edited by A. Bisin and G. Federico, 463–89. Amsterdam: Elsevier.

Prummer, A., and J.-P. Siedlarek. 2017. "Community Leaders and the Preservation of Cultural Traits." *J. Econ. Theory* 168:143–76.

Przeworski, A. 2004. "The Last Instance: Are Institutions the Primary Cause of Economic Development?" *European J. Sociology* 45 (2): 165–88.

Putnam, R. D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton Univ. Press.

Schiavone, A. (1996) 2020. *La Storia Spezzata: Roma Antica e Occidente Moderno*. Torino: Einaudi. Original edition: Bari: Laterza.

Swidler, A. 1986. "Culture in Action: Symbols and Strategies." *American Sociological Rev.* 51:273–86.

Tabellini, G. 2008a. "Institutions and Culture." *J. European Econ. Assoc.* 6 (2–3): 255–94.

———. 2008b. "The Scope of Cooperation: Norms and Incentives." *Q.J.E.* 123 (3): 905–50.

———. 2010. "Culture and Institutions: Economic Development in the Regions of Europe." *J. European Econ. Assoc.* 8 (4): 677–716.

Tadei, F. 2018. "The Long-Term Effects of Extractive Institutions: Evidence from Trade Policies in Colonial French Africa." *Econ. Hist. Developing Regions* 33 (3): 183–208.

Ticchi, D., T. Verdier, and A. Vindigni. 2013. "Democracy, Dictatorship and the Cultural Transmission of Political Values." IZA Discussion Paper no. 7441, Inst. Labor Econ., Bonn.

Touré, N. 2021. "Culture, Institutions and the Industrialization Process." *J. Econ. Behavior and Org.* 186:481–503.

Treisman, D. 2017. "Democracy by Mistake." Working Paper no. 23944, NBER, Cambridge, MA.

Valencia, F. 2021. "Historical Econometrics: Instrumental Variables and Regression Discontinuity Designs." In *Handbook of Historical Economics*, edited by A. Bisin and Federico, 179–211. Amsterdam: Elsevier.

Verdier, T., and Y. Zenou. 2015. "The Role of Cultural Leaders in the Transmission of Preferences." *Econ. Letters* 136:158–61.

———. 2018. "Cultural Leader and the Dynamics of Assimilation." *J. Econ. Theory* 175:373–414.

Volterra, V. 1926. "Variazioni e fluttuazioni del numero d'individui in specie animali conviventi." *Memorie dell'Accademia dei Lincei* 2:31–113.

Voth, H. J. 2021. "Persistence—Myth and Mystery." In *Handbook of Historical Economics*, edited by A. Bisin and G. Federico, 243–67. Amsterdam: Elsevier.

Weibull, J. W. 1995. *Evolutionary Game Theory*. Cambridge, MA: MIT Press.

Wittman, D. 1989. "Why Democracies Produce Efficient Results." *J.P.E.* 97 (6): 1395–424.