

Buchen, Clemens

Article — Published Version

Institutional reform paths

Economics & Politics

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Buchen, Clemens (2023) : Institutional reform paths, Economics & Politics, ISSN 1468-0343, Wiley, Hoboken, NJ, Vol. 35, Iss. 3, pp. 1099-1121, <https://doi.org/10.1111/ecpo.12252>

This Version is available at:

<https://hdl.handle.net/10419/288061>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>

Institutional reform paths

Clemens Buchen 

Economics Group, WHU—Otto
Beisheim School of Management,
Vallendar, Germany

Correspondence

Clemens Buchen, WHU—Otto Beisheim
School of Management, Vallendar,
Germany.

Email: clemens.buchen@whu.edu

Abstract

This paper develops a dynamic population game in which agents play a simple anonymous-exchange game of cooperating or defecting. Agents switch to the strategy with a higher expected payoff. Reformers can affect the payoff structure of the stage game to maximize the number of cooperators in the population by either enacting legal reform (institutional quality of contract law) or focusing on the macro outlook of the economy. Based on the theoretical model, the paper predicts which types the reformer should enact first and under which conditions reform will not be successful.

KEYWORDS

macroeconomic political economy, microfoundations of political economy, property rights, reform/stabilization, social choice

1 | INTRODUCTION

Generally speaking, a country's policymakers choose between two reform types, addressing institutional quality and macroeconomic stability. If reformers are constrained in terms of how much reform can be enacted in a given period, sequencing becomes relevant: which type of reform should be done first? I propose a model of reform sequence based on a population game. Agents in a population play a simple anonymous exchange game in which they can either carry out the promise of a contract or renege on it. Over time, agents adjust their behavior in the game, switching to the strategy with a higher expected payoff. A reformer can step in and alter the payoffs in the stage game, thereby influencing individual behavior, the overall payoff of the population, and the dynamic trajectory of the structure of the population.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Economics & Politics* published by John Wiley & Sons Ltd.

The model captures stylized facts about the interplay between fundamental coordination problems and a society's responses to address them. First, reform takes time because institutions can only be changed piecemeal. A reformer cannot simply choose an institution's "efficient" level but has to deal with initial conditions, which can be gradually adjusted. Second, different reform efforts can either aim at the general macroeconomic outlook of the economy or address a country's institutions. This means that distinct types of reforms can be implemented. Third, the optimal sequencing of the kinds of reforms depends on the fundamentals of the game agents play and the initial conditions a country faces. Fourth, reform can be too little, too late. Depending on the underlying structure of the societal game and the initial conditions, if any reform is enacted too late, the society will take a long time to reach the desired equilibrium because the population has already gone too far in the direction of the undesired equilibrium. On the other hand, it is also possible that late reform still leads to a success story. It depends on the game's structure, and the model can predict the conditions under which this is possible. Fifth, the same equilibrium play can be reached with different levels of institutional quality. This depends again on initial conditions and the type of chosen reform paths.

The basic game presents the fundamental challenge of anonymous exchange: will the contract partner fulfill the contract or not? Thus, any player employs one of two strategies, cooperate or deviate from the agreement. On this basis, reform will change the underlying game's payoff structure in two ways. First, strengthening the legal system, or, more precisely, institutional reform of contract law, makes it easier for a wronged party to seek damages in court. Second, the gains from cooperation can be increased by affecting the overall macro outlook of the economy. This distinction echoes Tinbergen's (1952) "qualitative" and "quantitative" reforms. The two reform effects constitute the main trade-off for a reformer. One result will be that, as a rule of thumb, improving the economic outlook is crucial if the gain from deviation is relatively high. In contrast, institutional reform becomes more critical the more significant the loss for the wronged party in the game. The game is then embedded into a population game, in which the dynamics of the different strategies can be studied using tools from evolutionary game theory. Individuals in the population change behavior according to which strategy promises higher expected payoffs. This creates different possible equilibria in the game, which can be stable or unstable. In this setup, the reformer's goal is to guide the population as fast as possible to a situation where all individuals stick to the agreed contracts. At the same time, the reformer considers the direct welfare effect any reform type will have. Reform, therefore, has a double impact. On the one hand, it influences behavior and, in this way, impacts the adjustment processes within a population. This constitutes an indirect effect on the total payoff from the population game. On the other hand, reform directly changes the total payoff in the population by affecting the payoffs of the underlying game. This dual effect plays also a crucial role in Dal Bó et al. (2018), who study voter support for policy changes that can transfer prisoners' dilemmas into cooperative games. They use lab experiments to show that voters tend to disregard the indirect effect.

There are reasons to believe that the baseline model underestimates the effects of the legal system because it focuses on the transfer of resources from a defendant to a plaintiff that comes from contract law. The legal system, of course, is broader than this. There is extensive literature discussing how to measure different aspects of the legal system (Gutmann & Voigt, 2018) and the importance of "good" legal rules for development, both from a theoretical (Torvik, 2020) and an empirical (Acemoglu et al., 2005) standpoint. Using one way to capture this, as an extension to the basic idea, I use the notion of *institutional resilience* (Buchen, 2022), which is related to a minimum level of the quality of the legal system. The concept of resilience is well established in ecology, where it refers to the ability of an ecological system to revert to its original form or function after a shock (Holling, 1973). The concept has made inroads into

economics mainly through dynamic risk management of environmental factors and climate change. See Li et al. (2017) for an overview of this literature. In the context of the model presented here, the notion of institutional resilience captures the role that a functioning legal system can play in averting an unraveling of the population in the wake of an exogenous shock. Given that the macro outlook in a population depends to a considerable degree on outside factors beyond a reformer's control, this variable can be affected by exogenous shocks. Technically, when an adverse shock hits, if the quality of the legal system is not sufficiently high, the population is at risk of being put on a path towards a stable no-cooperation equilibrium, from which it is hard to get out of. In this case, the exogenous shock can lead to a persistent crisis. A minimum level in the quality of the legal system allows the population to bounce back after a shock. It ensures that the adverse effects of an exogenous shock can be contained by stabilizing the population at a stable (interior) equilibrium and avoiding entering the basin of attraction of the no-cooperation equilibrium. This line of reasoning introduces a long-term consideration for a functioning legal system alongside the more immediate problem of guiding a population of agents toward cooperating.

The following section gives a brief overview of some additional literature. Then I introduce, solve and discuss the main model. Section 4 offers an extension along institutional resilience, and the last section provides a discussion and some concluding remarks.

2 | LITERATURE

While the model is meant to be abstract enough to capture generic reform efforts, it is inspired by the experience of postsocialist transition countries in the 1990s. Tracking their reform progress over time, a clear pattern emerges. While all Central and Eastern European transition countries fell into an early recession, after a while, some started growing consistently while others did not. The transition literature early on discussed the problem of sequencing of reform as one of the root causes for the divergent experiences of countries, that is, the fundamental question of which kind of institutional reforms should come first or whether reformers should first get the macroeconomic environment under control. There the focus was primarily on privatization and market liberalization efforts. Roland (2000), Campos and Coricelli (2002), and Havrylyshyn (2006) provide overviews of these debates. Relatedly, in his account of the transition period, Havrylyshyn (2006) develops a model in which early reforms have a lasting impact on a country because they set it on a path of a virtuous cycle, from which investment and growth follow. If early reforms do not come forth, corruption becomes entrenched, and the country is trapped in a hard-to-escape vicious cycle. My model captures this basic idea of a country slipping towards a bad equilibrium, although corruption per se is not needed for that, just plain inaction by policymakers and lawgivers. My approach is similar in spirit also to Ialnazov and Nenovsky (2011), who model the transition process as a game in which the equilibria change because of reform and economic development. They argue that the early period in the transition from socialism is characterized by the low-payoff equilibrium of a Prisoners' dilemma. Over time, payoffs and incentives change, making it possible for players to cooperate. In my model, I am making precise how the game changes.

The transition phase generally coincided with a heightened interest in questions of policy and reform. On the one hand, similar to the Central and Eastern European experience, earlier reforms in Latin America became a focus of analysis (Edwards, 1990). In the Latin American context, Martinelli and Tommasi (1997) pick up the question of big-bang versus gradualist

reforms. They develop a game-theoretic framework to understand when big-bang reforms are preferable to overcome a large number of veto players and uncertainty about how reform affects the payoff of different constituencies. Recently, Asturias et al. (2016) return to questions of sequencing and study specifically how different sequences of reform of entry costs, trade costs, and the enforcement of contracts can impact the distribution of firms. On the other hand, in addition to the reform sequence, a complementary question is why Pareto-improving reforms sometimes do not occur. One explanation put forth concerns the power of interest groups to block reform, which would hurt their group, at least temporarily. For example, Hoff and Stiglitz (2008) argue that an improvement of the rule of law in a country is opposed by those who benefited from the former “lawless” state. Any attempt by lobby groups to protect their rents can be an obstacle to the goal of moving away from an underdeveloped state (Rajan, 2009). For these and other reasons, populations can also end up in “institutional traps” (Polterovich, 2008) of Pareto-dominated institutional configurations. The model I am proposing contributes to this literature. One result is that an institutional trap of low cooperation can emerge without recourse to political-economy considerations. It is enough that reforms are not introduced quickly enough. In this scenario, the level of cooperative behavior in the population has fallen too low a level for individual incentives to skew towards switching to cooperation. This incentive-driven version of institutional traps is also present in Belloc and Bowles (2013).

In terms of modeling strategy, my approach is squarely in the tradition of Aoki (2001) and Bowles (2004). The basic idea is to express the coordination problems in all societies as simple stage games, which provide a micro foundation of interactions between firms, individuals, or other entities in an economy. A similar approach can also be found in the literature on governance mechanisms. For example, community enforcement (Ellison, 1994; Greif, 1993; Kandori, 1992) analyzes how the individual player’s incentive to deviate in strategic interactions is overridden by informal punishment by the local community. Relatedly, intermediaries can take over the role of enforcers (Dixit, 2003), for example, in the case of the law merchants in the middle ages discussed in Milgrom et al. (1990). The fundamental question boils down to how cooperation can be started and how it can be made to last (Carmichael & MacLeod, 1997; Kranton, 1996) and how formal and informal governance mechanisms can interact (Sobel, 2006). More broadly, in the direction of political economy, the basic approach in terms of methodology is also present in Nunn (2007) for the choice between productive work and crime, or Acemoglu and Jackson (2015) for general choices of cooperating or not cooperating, or, more recently, Bowles and Choi (2020) for the choice between dividing gains or fighting over them and its impact on the formation of property rights.

3 | A DYNAMIC MODEL OF REFORM

The fundamental idea of the model is the following. In an infinitely large population of agents, each agent is randomly matched to play a symmetric one-shot stage game with one other player. Therefore, in the first step, I will present an augmented version of a model used by Bowles (2004, chapter 7) as the stage game. It is essential to understand, and it will be made more precise at a later stage, that the agents adopt one of the pure strategies of this game as their employed strategy. As time goes by, agents will change strategies depending on the difference in expected payoffs between strategies. Based on the stage game, I will then discuss different reform types, how they impact the stage game, and how reform can develop over time. At that point, I will embed the stage game into a population game and describe the dynamics of

the prevalence of the strategies more precisely. Finally, I will formulate a dynamic optimization problem and present and discuss the results. As theoretical robustness checks, I will discuss the sensitivity of the results to some of the assumptions made during the modeling process.

3.1 | Stage game

The stage game has two possible strategies, *Cooperate* or *Defect*. This setup is meant to be general enough to capture all sorts of contractual relationships, which can be characterized as anonymous exchange. Suppose a contract is written for the production and delivery of a good, and this contract is not a spot contract. Suppose that contract partners have to make some intermediate investment to fulfill their end of the contract. This opens up the opportunity of expropriating a rent if reneging on the contract. In terms of this model, to play the strategy “Cooperate” is to honor the contract, and to play “Defect” is to renege on it, capturing a rent.

If both players cooperate, the pairing results in a payoff of 1 for each. Cooperation now allows for the possibility that agents parlay their random encounters into repeated interactions. With ρ as the continuation probability, the expected number of rounds becomes $\frac{1}{1-\rho}$. However, if the other player defects, a cooperator incurs a loss $b < 0$, and the relationship ends immediately. In this case, the cooperator becomes the plaintiff and seeks damages through the legal system. Damages in this model cover the difference between the expectation to earn 1 and the actual payoff b , so $1 - b$.¹ The variable $\theta \in [0, 1]$ captures the probability that the court awards damages. A defector can earn a rent of $c > 1$ if the partner happens to cooperate. If deemed liable by the court, the defector (the defendant) pays damages of $1 - b$. Again, this occurs with probability θ . If both defect, each earns 0, and the relationship ends immediately. Figure 1 gives the row player payoffs of the symmetric game.

I assume that $c > 1$ and $b + c < 2$. This implies that with $\theta = \rho = 0$, the game constitutes a Prisoners' dilemma with defecting as the dominant strategy, but in which cooperation of both players would be welfare maximizing. In a state in which players do not consider any continuation of cooperation possible and in which they would not have any way of recovering damages, the only equilibrium play is to defect.

3.2 | Reform types

Against the backdrop of the game agents in this society play, an institutional reformer can carry out two types of reform. The first type addresses institutional quality, which will be represented by the efficiency of the contract law. More to the point, efficiency is captured by θ , the probability that one contracting party can recover damages in a court of law. As mentioned

	Cooperate	Defect
Cooperate	$\frac{1}{1-\rho}$	$b + \theta(1 - b)$
Defect	$c - \theta(1 - b)$	0

FIGURE 1 Row player's payoffs in the stage game.

above, if Player 1 plays Cooperate while Player 2 plays Defect, Player 1 can seek redress in court. The higher the institutional quality of the legal system, the higher the probability that Player 1 is successful in gaining damages. This implies a broad range of what the reformer could specifically do to affect the variable θ . For example, a commercial code can be introduced, updated, or streamlined to make its application more readily available. Also, the training of judges can be improved, more personnel helping with a backlog of cases hired, or incentives for judges for a speedy resolution of cases as a whole can be made more favorable. All of this impacts how the institution of resolving a contractual dispute in courts can be improved and is all captured in the summary variable θ .

The second type of reform impacts the economic outlook of the economy in general and, in particular, how this influences players' evaluation of the likelihood of the continuation probability. This entails any measures stabilizing the macro outlook in the economy, for example, getting inflation under control. Recall from the discussion of the stage game that, in principle, the one-shot game can be transferred into a repeated relationship. This is captured by ρ , the probability that the pairing continues for one more round in the future. The more stable and predictable the economic outlook, the higher the likelihood that agents are planning ahead, and the (subjective) probability of repeated interactions increases.² If expectations about the future are favorable and predictable, that is, if agents expect future economic growth, low inflation, in general, a stable economic policy, and both agents play "Cooperate," this increases the chance that the one-shot relationship is developed into a recurring one.³ Keep in mind that the stage game attempts to capture anonymous exchange. It is undoubtedly the case that in periods of an unfavorable economic outlook, such as during a period of hyperinflation, individuals retreat to personal relationships, informal dealings, and barter. In terms of the stage game, this would be expressed by playing "Defect" in an anonymous setting. In my model, the result of the cooperation is the potential creation of *new* repeated business interaction, possibly taking up a lengthy project which would need continuous attention by both parties over an extended period. In this case, a favorable economic outlook fosters this kind of relationship. But one can also envision other types of reforms having a similar impact on the continuation probability, such as investment in infrastructure or entering into trade agreements. What is essential for the model at hand is that the reformer can favorably influence the probability of repeated interactions by creating a stable, predictable, and favorable economic outlook for the population. Stated differently, this reform variable aims to reduce uncertainty from the players' point of view. For the sake of the workings of the model, I assume that the reformer can directly impact the variable ρ .⁴

To summarize, I have two types of reform in mind. The first one increases the efficiency of the institutional set-up, specifically of the contract law. In this way, the off-diagonal payoffs of the game in Figure 1 are affected by decreasing the expected value of defecting and increasing, all else equal, the expected value of carrying out the contract. The second type of reform aims at making cooperation between individuals more attractive, independent of the institutional framework in which they find themselves. Technically, this affects the upper left cell of the game in Figure 1. The way the two types of reforms were defined allows me to distinguish their impact on the different payoffs in the cells of the payoff matrix.

Reform choice will be captured by the reform variable $\mu \in [0, 1]$. This can be interpreted as the weight the reformer puts on either of the reforms. This also implies that the reformer is constrained in terms of the total reform effort per period. Using the common dot notation for a

derivative with respect to time, the change of the two variables follows two differential equations:

$$\begin{aligned}\dot{\theta} &= \mu(1 - \theta), \\ \dot{\rho} &= (1 - \mu)(1 - \rho).\end{aligned}$$

The choice of the dynamics for θ and ρ is meant to capture the assumption that both reform types have a decreasing marginal product throughout the reform period. It also ensures that both variables stay in the unit interval. Focusing entirely on institutional reform (θ) means setting $\mu = 1$, whereas throwing all reform effort at the economic outlook (ρ) implies $\mu = 0$, accordingly.

The economy starts with initial values, θ_0 and ρ_0 . Initial conditions are crucial in this model. The reformer cannot simply choose an optimal level in either of the variables but instead has to decide which variable should be incrementally increased in each period. In other words, reform is piecemeal and cumulative.⁵ The reform variable and either reform type are costless for the reformer.

3.3 | Population dynamics

The next step is to embed the stage game into a dynamic population game. In each instant of time, members of the population are randomly paired to play the one-shot stage game of anonymous exchange. Each player in the population starts with a pure strategy they stick to. Denote v_i , $i \in \{c, d\}$ the expected values of each pure strategy, and x the fraction of the population employing strategy “Cooperate” (c) at any instant of time. Then $1 - x$ gives the fraction playing “Defect” (d). It is easiest to think of members of the population belonging to either of two types: cooperators or defectors, meaning they play either of the two pure strategies when they are matched to play the game. Over time, however, these roles can be changed as players learn about the relative payoffs from either of the two roles. In general, one would expect that over time those strategies that promise a higher expected payoff will increase in prevalence in the population, while those with a lower payoff should wane, or, formally:

$$\begin{aligned}\dot{x} > 0 &\Leftrightarrow v_c > v_d, \\ \dot{x} < 0 &\Leftrightarrow v_c < v_d, \\ \dot{x} = 0 &\Leftrightarrow v_c = v_d.\end{aligned}$$

Using a linear equation is the simplest way to formalize this:

$$\dot{x} = v_c - v_d. \tag{1}$$

The linear dynamics go back to Friedman (1991) where the right-hand side is still divided by the number of available strategies. Lahkar and Sandholm (2008) offer a behavioral interpretation: If a player experiences their chosen strategy earning less than an unweighted average of all available strategies, this player considers switching strategies. The propensity to

switch is greater the smaller the fraction of players still employing this player's strategy. For the game in Figure 1 the dynamic (1) becomes:

$$\dot{x} = \begin{cases} 0 & \text{if } x = 0, \\ x \left(\frac{1}{1-\rho} - b - c \right) + b + \theta(1-b) & \text{if } x \in (0, 1), \\ 0 & \text{if } x = 1. \end{cases} \quad (2)$$

This first-order differential equation has possibly one interior stationary point where $\dot{x} = 0$ for an $x \in (0, 1)$. For $x = 0$ and $x = 1$, stationarity has to be imposed; otherwise, the system would leave the simplex.

As seen from (1), the evolutionary process is analyzed on a deterministic model. Recall that, first, the number of individuals playing the stage game is assumed to be infinitely large, and second, the time horizon of the analysis is finite. If these two assumptions are met, the stochastic Markov process of the change of x over time can be approximated by the deterministic dynamic (1).⁶ The latter assumption of a finite time horizon is justified since the focus is on a period of institutional reform. The assumption of infinite population size is congruent with the aim to analyze the role of institutions in fostering *anonymous* exchange. If populations were small, the idea that contract partners are randomly matched to play an anonymous game would be lost.⁷

In general, depending on the parameter values, four game types can emerge; apart from the prisoners' dilemma (*PD*), a hawk-dove (chicken) game (*HD*), a coordination game (*CD*), and a cooperate-dominance game (*D*) are possible. The four game types and the stability of their rest points ($\dot{x} = 0$) are summarized in Table 1. In principle, each type can have up to three rest points; for example, the *HD* game has unstable rest points at the corners, $x = 0$ and $x = 1$, which implies that the monomorphic states in which only one type exists are unstable. A small perturbation in x would move the population away from either equilibrium. In the case of *HD*, the interior rest point exists and is stable. If the equilibrium were to be perturbed, the population would self-correct and move back toward the stable rest point. The interpretation is similar for the three remaining possible types. The notation introduced in Table 1 (last column) symbolizes this using arrows pointing away from (towards) the unstable (stable) rest point(s). This notation is carried over into Figure 2, which marks off the emerging game types in the θ - ρ -plane.

The problem of the reformer is to choose a path from an initial condition represented by a point in the plane in Figure 2. Depending on this choice, the population would go through different game types during the reform period.

TABLE 1 Stability of rest points in the four-game types.

Type	$x = 1$	$x = 0$	$x \in (0, 1)$	Notation
<i>PD</i>	Unstable	Stable	\nexists	\leftarrow
<i>HD</i>	Unstable	Unstable	Stable	$\rightarrow\leftarrow$
<i>CD</i>	Stable	Stable	Unstable	$\leftarrow\rightarrow$
<i>D</i>	Stable	Unstable	\nexists	\rightarrow

Abbreviations: CD, coordination game; D, cooperate-dominance game; HD, hawk-dove (chicken) game; PD, prisoners' dilemma.

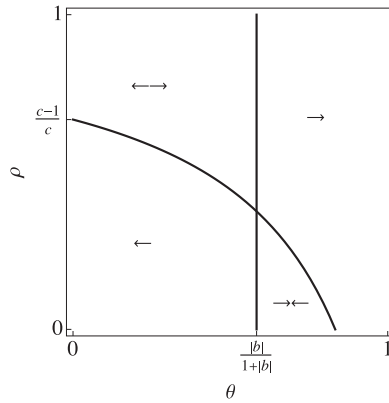


FIGURE 2 Game types.

3.4 | A dynamic optimization problem

The reformer's goal is to get the percentage of cooperators 100% as fast as possible by choosing the reform variable μ . In mathematical terms, a dynamic optimization problem needs to be set up, in which $x = 1$ is reached in some time $t = T$. To focus on the question of sequence, during the reform period, the question of welfare in the population is tuned out. However, the reformer realizes that once the goal is reached, this creates a payoff for the population. This implies that, in mathematical terms, the dynamic optimization problem involves a salvage value once the reform effort is over. The salvage value is given by the total payoff generated from the game $\Omega = \delta \left(2x^2 \left(\frac{1}{1-\rho} - b - c \right) + 2x(b + c) \right)$ with discount factor $\delta = \int_T^\infty e^{-rt} dt$, where r is the rate of time preference, for example, simply an interest rate. The salvage value is evaluated at $t = T$. At the time T , the reform effort is over, all variables become constants, and the payoff from the game accrues from then on. This is expressed by the discount factor δ .

Taken together, this set-up allows for a tractable model in which the two goals are still present: speed of reform and welfare of the population. The fact that welfare only explicitly enters at the end of the reform process does not entirely remove it from the trade-off between speed and welfare facing the reformer. As will become clear, although the salvage value only accrues later, it already impacts the reformer's choice during the reform phase.

Suppressing the dependence on the time for the variables x, ρ, θ , and μ , the dynamic optimization problem can be written as:

$$\min_{\mu} T - \Omega \quad \text{subject to} \tag{3}$$

$$\dot{x} = x \left(\frac{1}{1-\rho} - b - c \right) + b + \theta(1 - b), \tag{4}$$

$$\dot{\theta} = \mu(1 - \theta), \tag{5}$$

$$\dot{\rho} = (1 - \mu)(1 - \rho), \tag{6}$$

$$0 \leq \mu \leq 1, \quad (7)$$

$$x(T) = 1. \quad (8)$$

In words, the goal is to choose a reform variable $\mu \in [0, 1]$ in each period t to minimize the time T it takes to get to $x = 1$. Given the shortest time, the reformer aims for the highest possible population payoff once the goal is reached. Equations (4–6) are the state equations. The appendix provides a full-fledged mathematical treatment of this dynamic optimization problem, which provides the proofs for all subsequent claims.

Any solution to a dynamic optimization problem describes a temporal path of the choice variable. The following lemma summarizes it for the choice of the reform variable μ .

Lemma 1. *In each period t , the reformer chooses the reform variable $\mu(t)$ according to:*

$$\mu(t) = \begin{cases} 1 & \text{if } \Delta(t) > 0, \\ 0 & \text{if } \Delta(t) \leq 0. \end{cases}$$

with

$$\Delta(t) = \int_t^T \lambda_1 \left((1 - \theta)(1 - b) - \frac{x}{1 - \rho} \right) d\tau - \frac{\delta}{1 - \rho_T}, \quad (9)$$

and λ_1 as the costate variable associated with state Equation (4); λ_1 is a function of $\rho(t)$, b , c , t and T ; ρ_T is the value of ρ in time T .

Equation (9) expresses the fundamental challenge facing the reformer. At any moment, any decision taken about how to allocate reform effort will cause ripple effects through the entire reform process since changed variables in one period change the future path of the system. This is why the choice in each period t is formally expressed as an integral from t to the final period T . The costate variable, λ_1 , represents the dynamic shadow price of the fraction of cooperators, x . The most useful way to think about it is the following: λ_1 measures by how much faster the goal of $x = 1$ is reached, that is, by how much T is reduced, if in any period x were marginally increased. In this sense, it represents how much the reformer values having a higher value of x at any point during the reform process. Then the integral in (9) measures the differential impact between a reform on θ and a reform on ρ on the evolution of x . Suppose that $\mu = 1$. Then the evolution of x changes by $\frac{\partial \dot{x}}{\partial \theta} \cdot \dot{\theta}$, which is equal to $(1 - \theta)(1 - b)$ in each period. Similarly, suppose $\mu = 0$. Then the impact on the change in x is given by $\frac{\partial \dot{x}}{\partial \rho} \cdot \dot{\rho}$, which equals $\frac{x}{1 - \rho}$. The difference between the two effects is valued in each period by the future value of its impact on the shortening of the reform period, λ_1 .

Whereas both reforms impact the evolution of the variable x , ρ also increases the salvage value Ω in (3) because it leads to an increase of the payoff from the game independent of the value of x . This effect is represented by the term $\frac{\delta}{1 - \rho_T}$ in (9). This is not true for θ because it constitutes a transfer of resources from defendants to plaintiffs, which cancels in the welfare calculus. On a fundamental level, the reformer trades off different effects of reforms, both on the speed of reform and the direct change in welfare.

The foregoing discussion leads to the first result in the following:

Lemma 2. *In the last period T , the reformer chooses $\mu = 0$.*

This result is straightforward to see. Once the system has arrived at $x = 1$, the reformer has no use for a yet increased value of θ since, at this point, only the variable ρ will impact the welfare going forward. Relatedly, to choose a reform on ρ only ($\mu = 0$), it is sufficient to have the marginal benefit per time of increasing ρ greater than for θ :

Proposition 1. *$x_0 \geq (1 - \rho_0)(1 - \theta_0)(1 - b)$ is a sufficient condition for choosing $\mu = 0 \ \forall t \in [0, T]$.*

There are two reasons for this. First, as stated above, ρ has the double impact of both affecting the evolution of x and increasing welfare. Second, the marginal effect increases over time. There are increasing marginal returns to the macroeconomic reform modeled in this case. If it is preferable to reform only the macro outlook at the beginning of the reform, it remains preferable throughout the entire time.

In contrast, if the reformer chooses to start with reforming θ only ($\mu = 1$) in $t = 0$, then there will be one switch to ρ throughout the reform period at a time t^* . This follows from Lemma 2. In this case, it is necessarily true that in $t = 0$, the marginal benefit from increasing θ is greater than for ρ . This is made precise in the following Lemma.

Lemma 3. *$x_0 < (1 - \rho_0)(1 - \theta_0)(1 - b)$ is a necessary condition to choose $\mu = 1 \ \forall t \in [0, t^*]$ with $0 \leq t^* < T$.*

However, this condition is not sufficient because, as stated already above, not only does the marginal benefit of θ on the increase in x have to be larger than the marginal benefit of ρ on x , but this has to be larger than the effect ρ has on the welfare. One can imagine a situation in which the reformer would choose to reform ρ even if the marginal effect of θ on the evolution of x were stronger. This is the case if the added payoff to the population game as a whole that comes from ρ outweighs the slower adjustment of x .

In sum, there are two regimes that a dynamically optimizing reformer will choose. They are shown in a stylized way in Figure 3. Under the first “institutions-first”-regime, the reformer

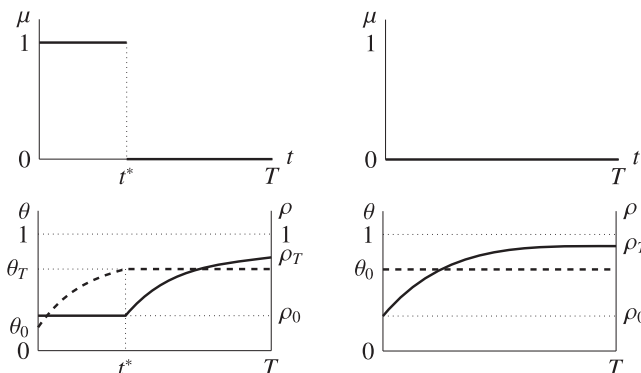


FIGURE 3 Two reform regimes.

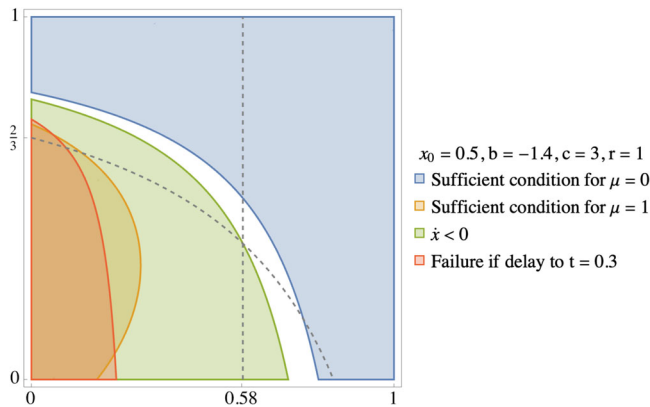


FIGURE 4 Numerical example. [Color figure can be viewed at wileyonlinelibrary.com]

will initially put all reform effort towards the legal system and then, at a time t^* , switch to improving—again, using all reform capacity—the economic outlook. This is shown in the left half of Figure 3. The upper panel shows the optimal choice of μ as a function of time t , and the lower panel depicts the resulting change of the variables θ (left vertical axis) and ρ (right vertical axis), also as a function of the time. The reform period is over when $x = 1$ is achieved in $t = T$. In the second regime (right-hand side of Figure 3), only the economic outlook is attended to for the entire time it takes the society to arrive at $x = 1$.

To illustrate further, Figure 4 shows the choice in $t = 0$ for different initial values θ_0 and ρ_0 . The initial situation has 50% of the population being cooperators in a game with $b = -1.4$ and $c = 3$. For all combinations of θ_0 and ρ_0 that lie within the green-colored region, the share of cooperators is declining ($\dot{x} < 0$). For all combinations within the yellow-shaded region, the reformer will embark on the institutions-first regime, whereas for all remaining possible combinations, the second regime of only reforming ρ for the entire reform period is preferred. Part of the latter—the blue-shaded area—has the marginal effect of ρ on x greater than for θ . The area between the yellow and the blue part represents all combinations of the initial values where ρ is chosen even though the direct effect coming from θ is larger. The dashed lines reproduce the delineations of game types from Figure 2 for this numerical example.

The following proposition summarizes some comparative statics results.

Proposition 2. *The reformer is more inclined towards reforming ρ only, (i) the greater x_0 , (ii) the greater θ_0 , (iii) the greater ρ_0 , (iv) the smaller $|b|$, and, (v) the greater c .*

Put simply, the worse the initial structural conditions, the higher the marginal payoff from increasing the institutional quality first. If things are bad, the reformer must focus on legal reform first, ignoring direct welfare effects. In terms of the exchange game that players play, the general guideline for the reformer should be, on the one hand, to make sure it pays to cooperate (impact on ρ) if the “temptation value” of defecting, c , is high, and, on the other hand, to make sure defecting is punished (increase θ) if the “sucker value,” $|b|$, is high.

There are conditions for which it is dangerous for a reformer to delay reform. Referring back to Figure 4, this will only be a problem if the population starts at a point where the change of cooperators is negative (green area). The following proposition provides a rule of thumb for a situation where even a prompt reform of any type would not lead to the desired outcome.

Proposition 3. *If, in any t , approximately $\ddot{x} \leq \frac{(\dot{x})^2}{2x}$, any reform regime will not be successful.*

The term \ddot{x} measures how the marginal product of any reform changes. The greater \ddot{x} , the more increasing the marginal returns of any reform on the evolution of x become over time. Stated differently, if these returns are (relatively) too small, reform will not carry enough heft to lift the population towards the goal of $x = 1$.

There is an important corollary to the statement expressed in Proposition 3. All else equal, the smaller \ddot{x} , the more dangerous it becomes for a reformer to delay reform because the population will have moved too far into the basin of attraction of $x = 0$. Going back to Figure 4, this is illustrated by the red-colored area. It represents the combinations of θ_0 and ρ_0 for which a delay of $t = 0.3$ for this number example would render any reform effort a failure.

The statement in Proposition 3 can be turned around and spelled out in more detail. Swift reform will be successful if, in any period, the following holds:

$$\dot{x} \left(\frac{1}{1-\rho} - b - c \right) + (1-\mu) \frac{x}{1-\rho} + \mu(1-\theta)(1-b) > \frac{(\dot{x})^2}{2x}. \quad (10)$$

This equation expresses the same fundamental idea discussed above but from a different vantage point. Suppose the reformer finds themselves in a situation where the mere success of bringing the population to $x = 1$ is in question, which is a situation where the hurdle of this inequality can only be cleared just so. In that case, the total focus must be on the relative effect of each reform on the evolution of x . These are given by the second and third terms on the left-hand side of (10), respectively. Then $(1-\theta)(1-b) > (<) \frac{x}{1-\rho}$ will lead the reformer to set $\mu = 1(0)$, because this maximizes the left-hand side of (10). As soon as the initial conditions are such that (10) is comfortably met, the trade-off between speed of adjustment and increasing the total payoff from the game comes into play. Stated differently, if there is no danger of reform failure, the reformer can afford to care about more than mere speed.

3.5 | Theoretical robustness checks

To conclude the model section, I return to some of the modeling choices and discuss whether and how they impact the results. This concerns the selection of the evolution of x in Equation (1), the objective function in Equation (3), and the decision not to include any cost of reform.

First, the choice of a linear dynamic in (1) is more general than it might appear at first glance. For two-strategy stage games of two players, the rest points and their stability are the same as in other commonly-used dynamics, such as the replicator dynamics (Taylor & Jonker, 1978), best-response dynamics (Hopkins, 1999), or dynamics based on differences in realized payoffs (Loginov, 2022). The out-of-equilibrium direction of change is also the same, but the speed of the linear dynamic differs from the alternatives mentioned above. In this sense, a linear choice for \dot{x} dramatically simplifies the analysis of the dynamics without substantially affecting any results.

Second, suppose that the objective function includes the average welfare of the population ($x^2/(1-\rho) + x(1-x)(b+c)$) instead of the speed of adjustment. In other words, the

reformer would still want to reach the state $x = 1$ but maximizes welfare during the reform path. All else equal, the following term would be added to the expression for $\Delta(t)$ in (9):

$$\int_t^T \frac{x^2}{1-\rho} d\tau.$$

This term captures the idea that any reform choice in a period t on ρ positively impacts the welfare immediately and not only at the end of the reform period. This is counteracted by the fact that a higher x also increases welfare. As a result, while this would tilt the choice towards choosing reform on ρ compared to the original model, the fundamental trade-off would not change. The reformer would still have to weigh speed against welfare. In this sense, the additional mathematical complication would not fundamentally change our understanding of the fundamental trade-offs.

Last, neither does the reform variable μ come with any costs attached, nor have I assumed any costs for tackling either of the two types of reform. Suppose one would go with the simplest way to include an increasing convex cost function for reform, for example, a function $C(\mu) = \frac{1}{2}\mu^2$. Then (9) would become $\mu = \Delta(t)$. As a result, rather than the sudden shift from $\mu = 1$ to $\mu = 0$ (“bang-bang”), $\mu(t)$ would have a continuous function allowing for periods in which both reforms would be enacted ($0 < \mu < 1$). This adjusted institutions-first regime is illustrated in Figure 5. From a qualitative point of view, nothing would change. In this case, it would still be true that one reform type would be prioritized over the other, effectively making $\mu = \frac{1}{2}$ the new focal point in which the reform effort swings from one type being predominant to the other at some time t' . The basic structure of the reform sequence would be unchanged. Similarly, I have not assumed specific costs for either type. One could imagine that one type, say θ , is more costly. This could tilt the balance in knife-edge cases but, again, would not affect the fundamental trade-off. Since any assumption about differential costs would necessarily be somewhat arbitrary, the no-cost assumption is the cleanest choice.

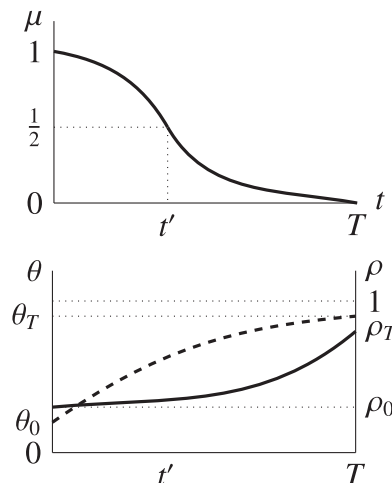


FIGURE 5 Reform regime with cost of reform.

4 | EXTENSION: INSTITUTIONAL RESILIENCE

As a rule, bad initial structural conditions will push a reformer to focus on institutions first. The model makes a strong case in favor of institutional reform as the basis of any reform effort. At the same time, the model arguably downplays the importance of institutional reform for at least two reasons. First, the model does not allow for spillover effects from institutions in the narrow legal sense to the overall institutional system of the economy and on the future macro outlook by individuals. More precisely, while the model allows for an indirect effect of institutions on welfare by influencing individuals' behavior, there might very well be a more direct effect from positive externalities that good institutions have on different social domains. This increases the importance of institutions even more.

In addition, institutional reform's importance is heightened for a second reason. Broadening the scope a bit, whereas the reformer has complete control over the quality of the legal institutions, the macroeconomic outlook, or, at least, how players perceive the level of uncertainty, will be influenced by factors beyond the reformer's control. Technically speaking, it is reasonable to assume that the value of ρ is partially exogenously determined. Take the simple example of a global economic slowdown due to conflict or a pandemic. In the language of the model presented here, this shock leads to a plummeting value of ρ . In this broader context, there is a case to be made that a minimum threshold for the quality of legal institutions must be reached to make the population resilient to an exogenous shock.

To appreciate the argument, the type of games the reformer can induce are central. See the discussion at the end of Section 3.3 and, in particular, Table 1 and Figure 2. For the following discussion, those game types are central in which the all-defectors rest point ($x = 0$) is unstable. This is true for the *HD* or *D* game types. The following Lemma introduces a minimum threshold for either of the two games to emerge:

Lemma 4. For $\theta \geq \frac{|b|}{1+|b|} \equiv \tilde{\theta}$, the game type is either *HD* ($\rightarrow\leftarrow$) or *D* (\rightarrow).

It follows from Lemma 4 that a population in which the quality of the legal system reaches at least a level of $\tilde{\theta}$, the level of cooperation becomes resilient to exogenous shocks on ρ ; for sufficiently high values of ρ , the population finds itself in a game in which cooperation is dominant (\rightarrow). If an exogenous shock decreases the value of ρ , the population moves into an *HD* game ($\rightarrow\leftarrow$). Denote \tilde{x} the stable interior rest point in an *HD* game. Either way, independent of the value of ρ , there can never be an unraveling of the level of cooperation beyond the point \tilde{x} because this rest point is stable. Making the institution of cooperation resilient (Buchen, 2022) to outside factors is then an additional focus next to the speed of adjustment.

Fundamentally, there are two ways to incorporate resilience into the model. First, introduce an additional constraint on the value of θ in the last period, $\theta_T \geq \tilde{\theta}$. Second, introduce a second objective function as payoff (salvage) once the reform process is over, which would measure the utility from having a value of θ beyond the threshold of $\tilde{\theta}$. The utility then comes from the fact that cooperation within the population is resilient to shocks. Capturing this in an objective function is not straightforward and open to interpretation. I propose the following option. Introduce an additional salvage function measuring utility *after* the reform process has been completed, $\omega(\tilde{x})$. This function is meant to measure the population game payoff from the safety net of a minimum level $\tilde{\theta}$ for institutional quality. If this level is achieved, the payoff from the

game will never fall below $\omega(\tilde{x})$. If both a constraint on θ_T and a salvage function $\omega(\tilde{x})$ are included in the optimization problem, $\Delta(t)$, the optimal decision presented in Lemma 1 must be restated:

Lemma 5. *With the additional constraint $\theta_T \geq \tilde{\theta}$ and the additional salvage function $\omega(\tilde{x})$, the optimal choice to minimize (3) through (7) is given by:*

$$\mu(t) = \begin{cases} 1 & \text{if } \tilde{\Delta}(t) > 0 \\ 0 & \text{if } \tilde{\Delta}(t) \leq 0 \end{cases}$$

with

$$\tilde{\Delta}(t) = \Delta(t) + \lambda_2(T)(1 - \theta_T) \quad (11)$$

In Equation (11), a positive term is added to the original expression for $\Delta(t)$. In the baseline model, the costate variable for Equation (5), λ_2 , is equal to zero in T . Crucially, $\lambda_2(T)$ now takes on a positive value. This implies that the scope for institutional reform (θ) increases because the reformer considers the long-term welfare effect of having θ at least equal to $\tilde{\theta}$. Similar to the discussion on the payoff from the game before, the goal of having a minimum level of θ by the end of the reform process already influences the choice during the reform. In particular, different from before, it is now possible to have $\mu = 1$ in the last period.

A high enough institutional quality can be understood as strengthening the institutional resilience of the population. Institutional resilience gained through the quality of the legal system neither prevents exogenous shocks from happening nor does it completely insulate the population from adverse effects on the payoff. But, in the wake of a shock, resilience can help stabilize the population and ensure that cooperation does not entirely break down.

5 | CONCLUSION

The paper sheds light on the dual impact of reforms. On the one hand, reforms change individual behavior because individuals respond to changed incentives. On the other hand, reforms directly impact the payoffs that players can gain from interaction with others. Both effects should be taken into consideration when thinking about reform sequences. It is true that a reformer focusing on speed will look at the differential impact of either reform type on the adjustment of the number of cooperators. At the same time, at the margin, it can be preferable to trade off the speed of reform against future payoff. The latter comprises the payoff from cooperation and the resilience that a functioning legal system provides.

The model captures the fundamental mechanisms driving reform efforts. It highlights the institutional complementarity between long-term cooperation and the strength of the enforcement of property rights. The legal system undergirds a population because it can help contain the damage in case of exogenous shocks. It is well known that long-term cooperation can be the result of self-enforcing, mutually beneficial contracts. But a functioning legal system prevents this system from entirely collapsing in the event of a crisis. This necessary input sustains a reform effort against adverse exogenous shocks over extended periods.

ACKNOWLEDGMENTS

Thanks to participants of the 2020 European Winter Meeting of the Econometric Society. Thanks to Benjamin Bental and Alberto Palermo for valued comments. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The author declares no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Clemens Buchen  <http://orcid.org/0000-0003-2448-2326>

ENDNOTES

- ¹ From a law-and-economics perspective, these are expectation damages. Other types exist, such as reliance damages – *b*. This would not change the analysis. See Hermalin et al. (2007) for an overview. I chose to model expectation damages for two reasons. First, according to Hermalin et al. (2007), they are most commonly applied in contract law. Second, in this model, they constitute the largest nominal value maximizing the expected damages.
- ² It should be stressed that ρ is not a rate of time preference or an interest rate even though it works similarly mathematically.
- ³ Inflation could easily be operationalized keeping ρ exogenous. With an inflation rate π future incomes are depreciated by a factor $(1 + \pi)^{-1}$, so the expected payoff from a repeated interaction would become $(1 - \rho(1 + \pi)^{-1})^{-1}$. Now, all else equal, decreasing inflation increases the expected payoff.
- ⁴ The later discussion will relax the assumption by adding the possibility that this probability is at least partially exogenously determined.
- ⁵ This is not to negate the fact that large reform efforts exist that would lead to large changes in either of the variables. The need to prioritize and choose really only arises if a reformer is constrained, therefore the focus on piecemeal reform.
- ⁶ This result is known as Kurtz's theorem (Kurtz, 1981), see also Sandholm (2011, p. 370).
- ⁷ In addition, the assumption of an infinitely large population allows me to ignore sub-population effects that come from those cooperators who form repeated interactions. These pairs are not taken out of the population but still play on to be potentially matched with a new partner where the probability of being matched with the same previous partner can be assumed to be zero.

REFERENCES

- Acemoglu, D., & Jackson, M. O. (2015). History, expectations, and leadership in the evolution of social norms. *The Review of Economic Studies*, 82(2), 423–456. <https://doi.org/10.1093/restud/rdu039>
- Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. In *Handbook of Economic Growth* (Vol. 1A, pp. 385–472). Elsevier. [https://doi.org/10.1016/S1574-0684\(05\)01006-3](https://doi.org/10.1016/S1574-0684(05)01006-3)
- Aoki, M. (2001). *Toward a comparative institutional analysis*. The MIT Press. <https://doi.org/10.7551/mitpress/6867.001.0001>
- Asturias, J., Hur, S., Kehoe, T. J., & Ruhl, K. J. (2016). The interaction and sequencing of policy reforms. *Journal of Economic Dynamics and Control*, 72, 45–66. <https://doi.org/10.1016/j.jedc.2016.05.006>
- Belloc, M., & Bowles, S. (2013). The persistence of inferior cultural-institutional conventions. *American Economic Review*, 103(3), 93–98. <https://doi.org/10.1257/aer.103.3.1>

- Bowles, S. (2004). *Microeconomics: Behavior, institutions, and evolution*. Oxford University Press. <https://doi.org/10.1515/9781400829316>
- Bowles, S., & Choi, J.-K. (2020). The neolithic agricultural revolution and the origins of private property. *Journal of Political Economy*, 127(5), 2186–2228. <https://doi.org/10.1086/701789>
- Buchen, C. (2022). Institutional resilience: How the formal legal system sustains informal cooperation. *Journal of Institutional Economics*, 1–15. <https://doi.org/10.1017/S1744137422000418>
- Campos, N. F., & Coricelli, F. (2002). Growth in transition: What we know, what we don't, and what we should. *Journal of Economic Literature*, 40, 793–836. <https://doi.org/10.1257/002205102760273797>
- Carmichael, H. L., & MacLeod, W. B. (1997). Gift giving and the evolution of cooperation. *International Economic Review*, 38(3), 485–509. <https://doi.org/10.2307/2527277>
- Dal Bó, E., Dal Bó, P., & Eyster, E. (2018). The demand for bad policy when voters underappreciate equilibrium effects. *The Review of Economic Studies*, 85(2), 964–998. <https://doi.org/10.1093/restud/rdx031>
- Dixit, A. (2003). On modes of economic governance. *Econometrica*, 71(2), 449–481. <https://doi.org/10.1111/1468-0262.00415>
- Edwards, S. (1990). The sequencing of economic reform: Analytical issues and lessons from Latin American experiences. *The World Economy*, 13(1), 1–14. <https://doi.org/10.1111/j.1467-9701.1990.tb00737.x>
- Ellison, G. (1994). Cooperation in the prisoner's dilemma with anonymous random matching. *The Review of Economic Studies*, 61(3), 567–588. <https://doi.org/10.2307/2297904>
- Friedman, D. (1991). Evolutionary games in economics. *Econometrica*, 59(3), 637–666. <https://doi.org/10.2307/2938222>
- Greif, A. (1993). Contract enforceability and economic institutions in early trade: The maghribi traders' coalition. *The American Economic Review*, 83(3), 525–548.
- Gutmann, J., & Voigt, S. (2018). The rule of law: Measurement and deep roots. *European Journal of Political Economy*, 54, 68–82. <https://doi.org/10.1016/j.ejpoleco.2018.04.001>
- Havrylyshyn, O. (2006). *Divergent paths in post-communist transformation: Capitalism for all or capitalism for the few?* Palgrave. <https://doi.org/10.1057/9780230502857>
- Hermalin, B. E., Katz, A. W., & Craswell, R. (2007). Contract law. In M. A. Polinsky, & S. Shavell (Eds.), *Handbook of Law and Economics* (pp. 3–138). Elsevier. [https://doi.org/10.1016/S1574-0730\(07\)01001-8](https://doi.org/10.1016/S1574-0730(07)01001-8)
- Hoff, K., & Stiglitz, J. E. (2008). Exiting a lawless state. *The Economic Journal*, 118, 1474–1497. <https://doi.org/10.1111/j.1468-0297.2008.02177.x>
- Holling, C. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4(1), 1–23. <https://doi.org/10.1146/annurev.es.04.110173.000245>
- Hopkins, E. (1999). A note on best response dynamics. *Games and Economic Behavior*, 29(1), 138–150. <https://doi.org/10.1006/game.1997.0636>
- Ialnazov, D., & Nenovsky, N. (2011). A game theory interpretation of the post-communist evolution. *Journal of Economic Issues*, 45(1), 41–56. <https://doi.org/10.2753/JEI0021-3624450103>
- Kandori, M. (1992). Social norms and community enforcement. *The Review of Economic Studies*, 59(1), 63–80. <https://doi.org/10.2307/2297925>
- Kranton, R. E. (1996). The formation of cooperative relationships. *The Journal of Law, Economics, and Organization*, 12(1), 214–233. <https://doi.org/10.1093/oxfordjournals.jleo.a023358>
- Kurtz, T. G. (1981). The central limit theorem for Markov chains. *The Annals of Probability*, 9(4), 557–560. <https://doi.org/10.1214/aop/1176994361>
- Lahkar, R., & Sandholm, W. H. (2008). The projection dynamic and the geometry of population games. *Games and Economic Behavior*, 64(2), 565–590. <https://doi.org/10.1016/j.geb.2008.02.002>
- Léonard, D., & Van Long, N. (1992). *Optimal control theory and static optimization in economics*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139173551>
- Li, C.-Z., Crépin, A.-S., & Folke, C. (2017). The economics of resilience. *International Review of Environmental and Resource Economics*, 11(4), 309–353. <https://doi.org/10.1561/101.00000096>
- Loginov, G. (2022). Ordinal imitative dynamics. *International Journal of Game Theory*, 51(2), 391–412. <https://doi.org/10.1007/s00182-021-00797-7>
- Martinelli, C., & Tommasi, M. (1997). Sequencing of economic reform in the presence of political constraints. *Economics and Politics*, 9(2), 115–131. <https://doi.org/10.1111/1468-0343.92024>

- Milgrom, P. R., North, D. C., & Weingast, B. R. (1990). The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs. *Economics and Politics*, 2(1), 1–23. <https://doi.org/10.1111/j.1468-0343.1990.tb00020.x>
- Nunn, N. (2007). Historical legacies: A model linking Africa's past to its current underdevelopment. *Journal of Development Economics*, 83(1), 157–175. <https://doi.org/10.1016/j.jdeveco.2005.12.003>
- Polterovich, V. (2008). Institutional trap. In S. N. Durlauf, & L. E. Blume (Eds.), *The New Palgrave dictionary of economics* (pp. 1–7). Palgrave MacMillan. https://doi.org/10.1057/978-1-349-95121-5_2717-1
- Rajan, R. G. (2009). Rent preservation and the persistence of underdevelopment. *American Economic Journal Macroeconomics*, 1(1), 178–218. <https://doi.org/10.1257/mac.1.1.178>
- Roland, G. (2000). *Transition and economics: Politics markets, and firms*. The MIT Press.
- Sandholm, W. H. (2011). *Population games and evolutionary dynamics*. MIT Press.
- Sobel, J. (2006). For better or forever: Formal versus informal enforcement. *Journal of Labor Economics*, 24(2), 271–297. <https://doi.org/10.1086/499973>
- Taylor, P. D., & Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1–2), 145–156. [https://doi.org/10.1016/0025-5564\(78\)90077-9](https://doi.org/10.1016/0025-5564(78)90077-9)
- Tinbergen, J. (1952). *On the theory of economic policy*. North-Holland Publishing Company.
- Torvik, R. (2020). Formal institutions and development in low-income countries: Positive and normative theory. In J.-M. Baland, F. Bourguignon, J.-P. Platteau, & T., Verdier (Eds.), *The handbook of economic development and institutions* (pp. 103–136). Princeton University Press. <https://doi.org/10.2307/j.ctvm7bbxr.7>

How to cite this article: Buchen, C. (2023). Institutional reform paths. *Economics & Politics*, 35, 1099–1121. <https://doi.org/10.1111/ecpo.12252>

APPENDIX

The Hamiltonian for the dynamic optimization problem (3) through (7) can be written as:

$$H = -1 + \lambda_1 \left(x \left(\frac{1}{1-\rho} - b - c \right) + b + \theta(1-b) \right) + \lambda_2 \mu(1-\theta) + \lambda_3(1-\mu)(1-\rho).$$

The salvage value is $\Omega = \delta(2x^2((1-\rho)^{-1} - b - c) + 2x(b+c))$. The necessary conditions are:

$$H_\mu = \lambda_2(1-\theta) - \lambda_3(1-\rho) \geq 0, \quad (\text{A1})$$

$$H_x = \lambda_1 \left(\frac{1}{1-\rho} - b - c \right) = -\dot{\lambda}_1, \quad (\text{A2})$$

$$H_\theta = \lambda_1(1-b) - \lambda_2\mu = -\dot{\lambda}_2, \quad (\text{A3})$$

$$H_\rho = \lambda_1 \frac{x}{(1-\rho)^2} - \lambda_3(1-\mu) = -\dot{\lambda}_3, \quad (\text{A4})$$

The transversality conditions are:

$$\lambda_1(T) = \frac{\partial \Omega}{\partial x} \Big|_{x=1} = \frac{4\delta}{1 - \rho_T} - \delta 2(b + c), \quad (\text{A5})$$

$$\lambda_2(T) = \frac{\partial \Omega}{\partial \theta} \Big|_{x=1} = 0, \quad (\text{A6})$$

$$\lambda_3(T) = \frac{\partial \Omega}{\partial \rho} \Big|_{x=1} = \frac{2\delta}{(1 - \rho_T)^2}. \quad (\text{A7})$$

Let $\Delta = \lambda_2(1 - \theta) - \lambda_3(1 - \rho)$. From (A1) it follows that $\mu = 1(0)$ if $\Delta > (<)0$. Differentiating Δ with respect to the time and using (5) and (6) yields:

$$\dot{\Delta} = (1 - \theta)(\dot{\lambda}_2 - \lambda_2\mu) - (1 - \rho)(\dot{\lambda}_3 - \lambda_3(1 - \mu)).$$

Substituting (A3) and (A4) and simplifying:

$$\dot{\Delta} = \lambda_1 \left(\frac{x}{1 - \rho} - (1 - \theta)(1 - b) \right).$$

Integrating yields

$$\Delta(t) = \int_t^T \lambda_1 \left((1 - \theta)(1 - b) - \frac{x}{1 - \rho} \right) d\tau - \frac{\delta}{1 - \rho_T}, \quad (\text{A8})$$

(Lemma 1).

This is equation (9) in the text. A few conclusions can be drawn by inspecting (9/A8). Clearly, $\Delta(T) < 0$, so $\mu = 0$ in the last period (Lemma 2). Suppose $\frac{x_0}{1 - \rho_0} \geq (1 - \theta_0)(1 - b)$. Then in $t = 0$ and any subsequent t , a negative number is added to the integral because the term $\frac{x}{1 - \rho}$ is increasing in ρ . Therefore $\mu = 0$ in all periods $t \in [0, T]$ (Proposition 1). This logic cannot be reversed; $\frac{x_0}{1 - \rho_0} < (1 - \theta_0)(1 - b)$ clearly is a necessary condition for $\mu = 1$ in $t = 0$, but it is not sufficient. Since $(1 - \theta)(1 - b)$ is decreasing and $\frac{x}{1 - \rho}$ is increasing in θ , the integral is neither necessarily positive nor is it necessarily greater than $\frac{\delta}{1 - \rho_T}$ (Lemma 3). The integral itself cannot be fully solved.

From (A1) the condition that $\mu = 1$ in $t = 0$ is

$$\lambda_2(0)(1 - \theta_0) > \lambda_3(0)(1 - \rho_0),$$

where λ_2 and λ_3 are evaluated at $t = 0$. Using a dynamic version of the envelope theorem (see Léonard and Van Long (1992, p. 153)), this can be rewritten as:

$$-\frac{\partial T}{\partial \theta_0}(1 - \theta_0) > \left(-\frac{\partial T}{\partial \rho_0} + \frac{\partial \Omega}{\partial \rho_0} \right) (1 - \rho_0). \tag{A9}$$

T is implicitly defined by:

$$e^{T\left(\frac{1}{1-\rho} - b - c\right)} \left(\frac{b + \theta(1 - b)}{\frac{1}{1-\rho} - b - c} + x_0 \right) - \frac{b + \theta(1 - b)}{\frac{1}{1-\rho} - b - c} = 1. \tag{A10}$$

Note that in (A10), both ρ and θ are also functions of T , and this equation does not allow for a closed-form solution of T . Instead, I use a quadratic approximation of the left-hand side of (A10) around $t = 0$. This approximated equation is given by:

$$x_0 + \dot{x}(0)T + \frac{1}{2}\ddot{x}(0)T^2 = 1. \tag{A11}$$

The quadratic approximation will only give a useful result for T in the case of $\ddot{x}(0) > 0$. In this case, there is only one positive solution to (A11):

$$T_1 = \frac{-\dot{x}(0) + \sqrt{(\dot{x}(0))^2 + 2\ddot{x}(0)(1 - x_0)}}{\ddot{x}(0)}.$$

Differentiating T_1 with respect to θ_0 and ρ_0 allows to state an approximated expression for (A9), and this expression was used to create the yellow-shaded area of the numerical example in Figure 2 in the text, where it was made sure that the values for $\lambda_2(0)$ and $\lambda_3(0)$ are positive in the relevant ranges. There is an additional transversality condition defining the optimal T : $H(T) + \partial\Omega/\partial T = 0$. But this condition is of little help in this case because it does not allow for a closed-form solution of the optimal T .

In the following case, any reform leads to $x = 0$:

$$x_0 + \dot{x}(0)T + \frac{1}{2}\ddot{x}(0)T^2 = 0. \tag{A12}$$

Here, the case $\ddot{x}(0) < 0$ does not need to be excluded; (A12) has two solutions if $0 < \dot{x}(0) \leq \frac{(\dot{x}(0))^2}{2x_0}$, but the following one is always the smaller one, which is also the only positive solution in the case of $\ddot{x}(0) < 0$:

$$T_0 = \frac{-\dot{x}(0) - \sqrt{(\dot{x}(0))^2 - 2\ddot{x}(0)x_0}}{\ddot{x}(0)},$$

(Proposition 3).

Most comparative statics results in Proposition 2 can be easily deduced from how the sufficient condition to set $\mu = 0$ changes (Proposition 1) in x_0, θ_0, ρ_0 and b . To arrive at some comparative statics results for c , first (A2) can be solved holding ρ constant and using (A5):

$$\lambda_1(t) = 2\delta \left(\frac{2}{1 - \rho_T} - b - c \right) e^{(T-t) \left(\frac{1}{1-\rho} - b - c \right)}.$$

Clearly, λ_1 is strictly decreasing in c . Next, a case distinction is needed. Suppose, first, that $\mu = 1$ in $t = 0$. Then (A3) and (A4) become:

$$\begin{aligned} \lambda_1(1 - b) - \lambda_2 &= -\dot{\lambda}_2, \\ \lambda_1 \frac{x}{(1 - \rho)^2} &= -\dot{\lambda}_3. \end{aligned}$$

Holding λ_1 , x , and ρ constant this system can be solved and evaluated in $t = 0$:

$$\begin{aligned} \lambda_2 &= (1 - b)\lambda_1 e^{-T}(e^T - 1), \\ \lambda_3 &= \frac{\lambda_1 T x}{(1 - \rho)^2} + \frac{1}{(1 - \rho_T)^2}. \end{aligned}$$

In this case, $\Delta(0) > 0$ (since $\mu = 1$) becomes:

$$\Delta(0) = \lambda_1 \underbrace{\left((1 - b)(1 - \theta)e^{-T}(e^T - 1) - \frac{T x}{1 - \rho} \right)}_{>0} - \frac{\delta(1 - \rho)}{(1 - \rho_T)^2} > 0.$$

Since the term in the parentheses must be positive to make $\Delta(0) > 0$, $\text{sign} \frac{\partial \Delta(0)}{\partial c} = \text{sign} \frac{\partial \lambda_1}{\partial c}$, and therefore $\frac{\partial \Delta(0)}{\partial c} < 0$.

The same steps for the case of $\mu = 0$ result in:

$$\Delta(0) = \lambda_1 \underbrace{\left(\frac{x}{1 - \rho} e^{-T}(e^T - 1) + T(1 - b)(1 - \theta) \right)}_{>0} - \frac{\delta(1 - \rho)}{e^T(1 - \rho_T)^2}.$$

The term in the parentheses is positive; therefore, in this case, it must also be true that $\text{sign} \frac{\partial \Delta(0)}{\partial c} = \text{sign} \frac{\partial \lambda_1}{\partial c}$ and therefore $\frac{\partial \Delta(0)}{\partial c} < 0$.

Game types

The four game types represented in Table 1 and Figure 2 can be derived directly from the definition of the games. Denote $p(i, j)$ the payoff from playing $i \in \{c, d\}$ if the opponent plays $j \in \{c, d\}$. The game classes follow from (with $p(d, d) = 0$):

$$\begin{aligned} PD: & p(c, c) \leq p(d, c) \text{ and } p(c, d) \leq 0, \\ HD: & p(c, c) \leq p(d, c) \text{ and } p(c, d) \geq 0, \\ CD: & p(c, c) \geq p(d, c) \text{ and } p(c, d) \leq 0, \\ D: & p(c, c) \geq p(d, c) \text{ and } p(c, d) \geq 0. \end{aligned}$$

The delineation of the areas in Figure 2 follows from these conditions. Note that an *HD* game only occurs if $b + c > 1$. For both *HD* and *D* games $p(c, d) \geq 0$, or, $b + \theta(1 - b) \geq 0$. Solving for θ gives the threshold stated in Lemma 4.

Institutional resilience

Introducing the condition $\theta_T \geq \tilde{\theta}$ and an additional salvage $\omega(\tilde{x})$ changes the transversality condition (A6). The new conditions read:

$$\lambda_2(T) - \frac{\partial \omega(\tilde{x})}{\partial \theta_T} \geq 0,$$

$$\left(\lambda_2(T) - \frac{\partial \omega(\tilde{x})}{\partial \theta_T} \right) (\theta_T - \tilde{\theta}) = 0.$$

Following the same steps that led to (A8) results in equation (11) in Lemma 5, which is repeated here:

$$\tilde{\Delta}(t) = \Delta(t) + \lambda_2(T)(1 - \theta_T).$$

In principle, both earlier reform paths are still possible. In addition, since now $\lambda_2(T) \geq \frac{\partial \omega}{\partial \theta_T} > 0$, it is possible to have $\mu = 1$ in $t = T$.