The Poor, the Rich and the Enforcer: 
Institutional Choice and Efficiency

Thorsten Koeppl 
Queen’s University

Cyril Monnet 
Federal Reserve Bank of Philadelphia

Erwan Quintin* 
Federal Reserve Bank of Dallas

October 20, 2008

Abstract

Are efficiency considerations important for understanding differences in the development of institutions? We model institutional quality as the degree to which obligations associated with exchanging capital can be enforced. Establishing a positive level of enforcement requires an aggregate investment of capital that is no longer available for production. When capital endowments are more unequally distributed, the bigger dispersion in marginal products makes it optimal to invest more resources in enforcement. The optimal allocation of the aggregate institutional cost is not monotonic in wealth and entails a redistribution of capital endowments before production begins. Investments in enforcement benefit primarily agents at the bottom of the endowment distribution, and lead to a reduction in consumption and income inequality. Efficiency, redistribution and the quality of institutions are thus intricately linked and cannot be studied independently.

Keywords: Enforcement; Institutions; Inequality; Human and Physical Capital.

JEL Classification: D31; D52; O11; O43.

*E-mail: thor@econ.queensu.ca, cyril.monnet@phil.frb.org, erwan.quintin@dal.frb.org. We would like to thank Russ Cooper, Per Krusell, Huw Lloyd-Ellis, Elias Papaioanou, B Ravikumar, Richard Rogerson and seminar and conference participants at the Federal Reserve Bank of Richmond, the University of Montreal, Rice University, UQAM, the University of Western Ontario, Ryerson University, Ente Enaudi, the SED meetings, the Midwest Theory Meetings and the Vienna Workshop in Macroeconomics for many helpful comments. The views expressed in this paper are not necessarily those of the Federal Reserve Bank of Philadelphia, the Federal Reserve Bank of Dallas or the Federal Reserve System.
1 Introduction

Well-functioning institutions play a pivotal role in economic development and growth. As Rajan and Zingales (2003) put it, proper institutions allow economies to “[unleash] the power of financial markets to create wealth and spread opportunity.” North (1990), Hurwicz (1994) and others have pointed out that in order to achieve an efficient allocation of resources, an economy requires institutions that can enforce obligations from trade as well as property rights. Somewhat surprisingly then, ineffective institutions appear to be remarkably persistent in many countries. What causes some countries to develop well-functioning institutions, while others, even at similar stages of economic development, don’t? More generally, what determines how institutions develop?

A large literature – see for instance, Persson and Tabellini (1994), Acemoglu and Johnson (2005), or Acemoglu and Robinson (2008) among many others – argues that institutional choices are largely driven by strategic or political considerations. While these factors are clearly relevant, it is puzzling that simple notions of efficiency have not been equally emphasized as potential explanations for how institutions develop. In this paper, we provide a tractable way to model institutional choice and explore whether efficiency considerations can generate differences in the quality of institutions across economies.1

The basic idea we develop in this paper is that building institutions is akin to the adoption of a new, better technology. Like new technologies, institutions are costly to set up, since they require the investment of resources that cannot be used for other productive purposes. Although investing in institutions is not directly productive, it enables a better allocation of resources – here, through enforcing the obligations from trade.

To make these ideas precise, we consider an economy where all agents operate the same production technology characterized by decreasing returns to scale. Agents are alike except in that some have a higher endowment of capital than others. Since returns to scale are decreasing, there are potential gains associated with redistributing capital across agents. However, agents will agree to a reallocation only if they can be compensated. This requires that a transfer from agents that receive capital can be enforced after production has taken place. We assume here that establishing such a positive level of enforcement requires an investment of capital. Then, a trade-off arises as capital can be used either as an input in production or to

---

1This is not to say that this literature has not recognized efficiency or, more generally speaking, economic considerations as important (see for example Acemoglu, Johnson and Robinson, 2005). This paper, however, places the emphasis on efficiency and does not take into account any political or strategic considerations.
increase enforcement. The benefits of enforcement depend on the potential gains from trade and hence on economic fundamentals, namely the initial discrepancy in capital endowments. As endowment inequality increases, benefits from enforcement rise.

Who should pay for setting up the institutions? Until enforcement is introduced, its financing cannot be imposed on agents and contributions have to be made on a voluntary basis. In other words, efficient institutions cannot be set up independently from the way they are financed. We provide a detailed characterization of the distribution of the institutional cost by showing how a market mechanism can achieve the efficient social arrangement. In this environment, the efficient investment in enforcement is financed by fees that grant agents access to the market. We show that agents' contributions to the institutional cost eventually rise with endowments, but that they are not monotonic in wealth. Efficiency considerations alone thus suggest that investing in enforcement should be accompanied by a redistribution of endowments, although the optimal redistribution does not take the form of a simple mean-preserving contraction as it would in environments where market imperfections are exogenous and costless. Nevertheless, we show that the income and consumption gains associated with the optimal social arrangement accrue primarily to agents at the bottom of the endowment distribution.

This points to an intricate, dynamic relationship between inequality, redistribution and institutional quality. Agents with poor capital endowments reap the benefits of better institutions, which enables them to accumulate more resources over time. Therefore, assuming simple intergenerational linkages, investment in enforcement reduces endowment and income inequality over time. As the dispersion in endowment diminishes, so do differences in marginal products across agents, and optimal investment in enforcement also declines. Differences in marginal products, however, never fall below a threshold that would make investment in enforcement unnecessary.

Our results have important implications for the growing literature on the effects of inequality on economic growth. This literature has established a theoretical link between endowment

Our set-up follows Koeppl (2007) who studies a similar trade-off in the context of intertemporal risk sharing.

Modeling enforcement as a technology can thus be interpreted in two complementary ways. Enforcement allows one to set up a market environment for decentralized exchange of capital. This includes the institutions necessary to enforce property rights and the obligations that arise from trade. Alternatively, the technology can be seen as a redistributive tax system directly, where capital is reallocated before production through taxation and where transfers redistribute income after production. In both interpretations, investment costs are incurred ex-ante to set up the necessary institutions.
inequality, the absence of good institutions – often modeled via an exogenous degree of market incompleteness – and economic development, suggesting that redistributing productive resources can foster growth.\textsuperscript{4} We show that once one treats market incompleteness as endogenous, the optimal policy should comprise investments in institutions that improve the functioning of markets. In fact, while the optimal solution also calls for some redistribution of endowments in our model, this redistribution is very different in nature from a simple mean-preserving contraction. Furthermore, institutional investments make that redistribution possible in the first place by allowing agents who contribute to it to benefit from trade.\textsuperscript{5}

The endogeneity of market incompleteness greatly complicates the interpretation of standard empirical tests of the predictions of models of inequality and growth. For instance, Perotti (1994) argues that conditional on “the degree of market imperfection”, inequality has a negative effect on aggregate investment in cross-country data. Likewise, after controlling parametrically for the initial level of inequality, proxies for the degree of market imperfection have a negative effect on investment. Even assuming that both wealth inequality and market imperfections are captured in a satisfactory fashion in these empirical exercises, the fact that the two variables have causal effects on each other makes quantifying their respective importance impossible outside of a structural model.

More basically, our work makes it clear that standard measures of inequality do not allow for adequate tests of models of inequality and growth.\textsuperscript{6} Our theory establishes a link between the dispersion of marginal products, economic development and institutional choice. A high degree of wealth inequality may or may not imply a large ex-ante dispersion in marginal products. We illustrate this critical point by formalizing a hypothesis recently formulated by Engerman and Sokoloff (2000). They surmise that initial inequality could explain why some nations in the western hemisphere were much faster to develop institutions conducive to trade than others. Engerman and Sokoloff (2002) describe XIXth century Latin America as an area where both human and physical capital were highly concentrated. The United States and Canada, however, had less inequality in both forms of capital, and developed

\textsuperscript{4}See Benabou (1996), Aghion (1998) or Aghion, Caroli and Garcia-Penalosa (1999) for a review of this argument.

\textsuperscript{5}Market allocations are often contrasted with redistribution through governments. See most recently Acemoglu, Golosov and Tsyvinski (2007) and the discussion of Albanesi (2007). To the contrary, we point out that investment in market-supporting institutions and redistribution should optimally complement each other.

\textsuperscript{6}This goes beyond the well-known issues associated with the scarcity and poor quality of wealth inequality data. Given these issues, Perotti (1994) uses income inequality as a proxy for wealth inequality.
better market-supporting institutions. We argue that if human and physical capital are complementary in production, a sufficiently high positive correlation between human and physical capital endowment leads to small differences in marginal products across agents, which reduces incentives to invest in institutions conducive to trade.\(^7\)

By focusing solely on efficiency considerations, our work complements the vast literature that studies political forces as the principal determinant of institutional development. Our paper suggests that in order to understand the relationship between institutions, inequality and economic performance, one must look at where the gains from such institutions arise and how these gains can be redistributed throughout the economy. If institutions persist to be bad due to political factors, this must come from the fact that one cannot allocate the benefits from institutions in such a way as to compensate those who pay for them.

2 The Model

We will first study the determinants of institutional quality in a static environment populated by a continuum of agents indexed by \(i \in [0, 1]\). Each agent is endowed with a quantity \(a_i \geq 0\) of capital, where we assume that \(a_i\) increases with \(i\) on \([0, 1]\). This distribution of endowments is non-degenerate so that there is some measurable inequality among agents, and is sufficiently smooth to allow us to use standard variational arguments.

Agents are also endowed with a technology that transforms an input \(k \geq 0\) of capital into a quantity \(k^\alpha\) of a consumption good where \(\alpha \in (0, 1)\). We adopt this Cobb-Douglas specification of the production schedule for concreteness. Assuming some form of decreasing returns to scale suffices, however, to derive our results. All agents seek to maximize their end-of-period income hence consumption of the single consumption good.

Since the individual production technology exhibits decreasing returns to scale and endowments differ across agents, marginal products also differ across agents. Therefore, agents have an incentive to trade capital before production starts: agents with low capital endowments would like to borrow some capital for production from other agents in exchange for post-production transfers of the consumption good. We assume, however, that enforcement is limited. Agents can default on any transfer they owe after production, in which case they incur a fixed cost \(\eta > 0\) denominated in consumption-equivalent units.

\(^7\)Doepke and Eisfeldt (2007) offer a different explanation for the Engerman-Sokoloff hypothesis that combines economic and strategic considerations.
This formulation of the default option follows Sappington (1983) and Banerjee and Newman (1993), among others. We interpret $\eta$ as an absconding cost. By paying this fixed cost, agents erase any payment obligation they face. Our results, however, can easily be generalized to a broader class of default cost specifications. One could assume for instance that default costs rise with the quantity of capital a given agent chooses to use in production (or, equivalently, with output) according to a schedule $D(\eta, k)$ for all $\eta, k \geq 0$. The tools we use in this paper would simply require that $D$ be jointly concave.

Our main goal is to endogenize the level of enforcement when improving institutions is costly. In this context, a natural trade-off arises between devoting resources to institutional building and to production. Studying this trade-off will enable us to characterize the determinants of institutional quality. We assume therefore that establishing an economy-wide enforcement level $\eta \geq 0$ requires an aggregate capital cost $g(\eta) \geq 0$ that must be borne before production begins. We also assume that the cost function $g$ is strictly convex, strictly increasing and twice differentiable on $(0, +\infty)$, and that $g(0) = 0$. To allow for fixed costs associated with a positive level of enforcement, we only require that $\lim_{\eta \downarrow 0} g(\eta) \geq 0$.

3 Institutional Choice and the Allocation of Capital

3.1 Capital Allocation with an Exogenous Level of Enforcement

As a benchmark for the subsequent analysis, we consider first the situation where the level of enforcement $\eta$ is given exogenously and does not require any investment of capital. In that context, all capital can be used in production of the consumption good and the exogenous level of enforcement determines what trades between agents can be supported.

Assume then that agents can trade capital before production takes place in a competitive market taking as given the gross interest rate $R > 0$ for renting capital.\footnote{The model then becomes a simplified version of the canonical set-up studied by much of the literature on inequality and growth. See for example Aghion (1998) and Benabou (1996).} Given the exogenous level of punishment, the agent maximizes his income taking into account the cost of renting additional capital

$$\max_{k_i} k_i^\alpha + (a_i - k_i)R$$

subject to

$$(k_i - a_i)R \leq \eta.$$
The constraint reflects the fact that the agent cannot commit to make a post-production payment that exceeds the cost of defaulting. The value of this constraint depends on the interest rate \( R \) which in equilibrium is such that the market for capital clears. The following result summarizes how equilibria depend on the exogenous level of enforcement.

**Proposition 3.1.** Given \( \eta \), a unique competitive equilibrium exists. Furthermore, aggregate output and the equilibrium interest rate \( R \) rise with enforcement level \( \eta \).

**Proof.** The aggregate supply of capital, \( \int a_i \, di \), is independent of both \( \eta \) and of \( R \). Given \( \eta \), increases in \( R \) lower capital use because this makes capital more costly and makes borrowing constraints tighter. Therefore, the demand schedule is monotonically decreasing in \( R \), implying that equilibria are unique.

For the second part of the proposition, fix \( \eta \) and let \( R(\eta) \) be the corresponding equilibrium price of capital. One easily shows that the equilibrium allocation of capital for a given \( \eta \) solves

\[
\max_{\{k_i \geq 0\mid i \in [0, 1]\}} \int k_i^\alpha \, di
\]
subject to:

\[
\int k_i \, di \leq \int a_i \, di,
\]

\[(k_i - a_i)R(\eta) \leq \eta \quad \text{for almost all } i.\]

Indeed, any solution \( \{k_i \geq 0\mid i \in [0, 1]\} \) to this constrained maximization problem must satisfy, for almost all \( i \),

\[
\alpha k_i^{\alpha - 1} - \theta - \rho_i R(\eta) = 0
\]
\[
\rho_i (\eta - (k_i - a_i)R(\eta)) = 0,
\]

for some non-negative Lagrange multipliers \( \{\theta, \rho_i \mid i \in [0, 1]\} \). The competitive allocation satisfies those conditions with \( \theta = R(\eta) \) and, for almost all \( i \), \( \rho_i = \frac{\alpha k_i^{\alpha - 1} - R(\eta)}{R(\eta)} \).

Inspection of this maximization problem shows that the desired result follows provided \( \frac{\eta}{R(\eta)} \) rises with \( \eta \). Assume to the contrary that \( \eta \) rises to \( \eta' \), but that \( \frac{\eta}{R(\eta)} > \frac{\eta'}{R(\eta')} \). Then all agents are more borrowing constrained and capital becomes more expensive. Hence, capital demand falls for all agents. But that contradicts the fact that both \( R(\eta) \) and \( R(\eta') \) are equilibrium rates, since aggregate capital supply is the same in both cases. \( \Box \)
Figure 1: Capital allocation ($\eta$ exogenous; no cost)

Figure 3.1 illustrates this result. For the given level of enforcement level $\eta$, the implied equilibrium interest rate $R(\eta)$ determines a unique optimal scale $k(\eta)$ of production. Agents with endowment past threshold $a(\eta) \equiv k(\eta) - \frac{\eta}{R(\eta)}$ are unconstrained. All other agents are constrained and operate at a lower scale with capital input given by the borrowing constraint, or $k_i = a_i + \frac{\eta}{R(\eta)}$. In equilibrium, capital input rises therefore one-for-one with initial endowments until agents become unconstrained.

It is clear that when $\eta$ is sufficiently high, no agent is constrained and markets are effectively complete. On the other hand, when $\eta = 0$, trade is impossible and autarky results. More generally, a higher enforcement level leads to higher interest rates and higher aggregate output. When enforcement increases from $\eta$ to $\eta'$, borrowing constraints are relaxed and the demand for capital shifts up at all rental rates. As the total supply of capital is fixed, it must be the case that the equilibrium interest rate $R$ increases. A higher equilibrium interest rate causes the optimal scale of production for unconstrained agents to fall. Because the total supply of capital is fixed, the two schedules must cross and it must therefore be the case that $\eta'/R(\eta') > \eta/R(\eta)$. As a result, as figure 3.1 makes clear, the dispersion in marginal products declines, and, correspondingly, output increases. Hence, improvements in enforcement raise
output by allowing marginal products of capital to become more homogenous through trade, much like a direct redistribution of endowments would.

3.2 The Optimal Level of Enforcement

The foregoing analysis suggests that economies where endowments are unevenly distributed have an incentive to improve enforcement to make it possible to reduce the dispersion of marginal products through trade. Such an improvement in institutional quality, however, presumably requires resources that are no longer available for production. This is now captured by the aggregate cost $g(\eta)$ of capital that is necessary to achieve $\eta$ level of enforcement.

To study the resulting trade-off between using endowments for institutional building and using them in production, we will first consider a planner whose objective is to maximize aggregate consumption. This focuses the analysis on pure efficiency considerations since such a planner has no direct interest in reducing inequality.\(^9\) We will also argue that the optimal allocation that emanates from this social planner problem can be supported as an equilibrium in decentralized markets.

Given a distribution of endowments, the planner proposes a capital allocation $k = \{k_i \geq 0 | i \in [0, 1]\}$, a schedule of post-production transfers $t = \{t_i \in \mathbb{R} | i \in [0, 1]\}$, and a level of enforcement, $\eta \geq 0$. The planner’s proposal is restricted in two ways. First, agents can choose to stay in autarky rather than participate in the proposed arrangement. Second, agents can decide to default on the post-production transfer stipulated by the planner in which case they incur the punishment of $\eta$ in equivalent units of consumption. Therefore, the planner solves:

$$\max_{(k,t,\eta)} \int (k_i^\alpha - t_i) \, di$$  \hspace{1cm} (3.3)

\(^9\)In addition, one can show that this entails no loss of generality. Assuming instead that the planner maximizes a strictly concave welfare functional over agents end-of-period consumption does not change the nature of the optimal allocation. The inclusion of ex-ante participation and ex-post enforcement constraints implies that there is no room for redistributing income at the end of the period once efficiency considerations have been taken into account.
subject to

\[
\int k_i di + g(\eta) = \int a_i di \tag{3.4}
\]
\[
k_i^\alpha - t_i \geq a_i^\alpha \quad \text{for almost all } i \tag{3.5}
\]
\[
t_i \leq \eta \quad \text{for almost all } i \tag{3.6}
\]
\[
\int t_i di \geq 0. \tag{3.7}
\]

The first constraint is a resource feasibility constraint. The second set of constraints stipulates that agents have to be willing to participate in the proposed arrangement given that they can always opt for autarky. The third set of conditions expresses the fact that the enforcement of transfers is limited. The final constraint states that the planner cannot invent resources at the end of the period: aggregate consumption cannot exceed aggregate output. Naturally, this constraint will bind at the optimal allocation, so that maximizing aggregate consumption is equivalent to maximizing aggregate output.

A generically unique solution to this problem exists, as we demonstrate in the appendix. We will now argue that because raising the enforcement level is costly, the planner never chooses to eliminate all inequality in marginal products.

**Proposition 3.2.** The optimal allocation with endogenous enforcement is such that \( k_i \) is not almost everywhere equal.

**Proof.** Denote the non-negative multipliers associated with the constraints (3.4) - (3.7) by \( \theta \), \( \{\lambda_i|i \in [0,1]\} \), \( \{\mu_i|i \in [0,1]\} \) and \( \tau \), respectively. If the planner chooses not to invest in enforcement (\( \eta = 0 \)), then \( k_i = a_i \) almost everywhere and the result holds trivially. Assume then that \( \eta > 0 \). Necessary conditions for an interior solution to the planner’s problem are given by\(^{10}\)

\[
\theta g'(\eta) - \int \mu_i di = 0 \tag{3.8}
\]
\[
\alpha k_i^\alpha - 1 (1 + \lambda_i) - \theta = 0 \quad \text{for almost all } i \tag{3.9}
\]
\[
-1 - \lambda_i - \mu_i + \tau = 0 \quad \text{for almost all } i. \tag{3.10}
\]

Also note that, together with the usual slackness conditions, these conditions are sufficient.

\(^{10}\)See Cesari (1983). Our set of constraints satisfies a standard constraint qualification.
Assume now, by way of contradiction, that \( k_i \) is constant a.e. Then, \( \lambda_i \) is constant a.e. as well by condition (3.9), as is then \( \mu_i \) by condition (3.10). But since \( \eta > 0 \), resource feasibility requires that \( a_i > k_i \) for a non-negligible set of \( i \). This implies that \( t_i < \eta \) and \( \mu_i = 0 \) for that set. Hence, we need \( \mu_i = 0 \) for almost all \( i \), which cannot be the case by condition (3.8), given that \( \eta > 0 \) and that, by (3.9), \( \theta > 0 \).

The intuition for this result is straightforward. When capital use is equated across agents, marginal products are equated as well. It follows that small deviations from such an allocation have a negligible impact on output. On the other hand, reducing enforcement has a first-order effect on the resources available for production. With this result in hand, we can now characterize the optimal allocation of capital more precisely.

**Proposition 3.3.** The optimal allocation of capital is characterized by two endowment thresholds \( 0 \leq a \leq \bar{a} \) and two bounds on capital use \( 0 \leq k \leq \bar{k} \) that determine the optimal capital allocation for almost all \( i \) according to

\[
   k_i = \begin{cases} 
   \frac{k}{1 + \lambda_i} & \text{if } a_i \leq a \\
   \frac{(a_i^\alpha + \eta)^{\frac{1}{\alpha}}}{\alpha} & \text{if } a_i \in [a, \bar{a}] \\
   \frac{\bar{k}}{1 + \mu_i} & \text{if } a_i \geq \bar{a},
   \end{cases}
\]

where \( \eta \) is the optimal level of enforcement.

**Proof.** As \( \eta = 0 \) implies autarky in which case our characterization holds trivially, we assume that the optimal solution has \( \eta > 0 \). Conditions (3.8)-(3.10) together with the associated slackness conditions describe the optimal solution. We first establish that for almost all agents, either the participation or the enforcement constraint holds with equality. Assume to the contrary that for some non-negligible set of agents constraint is not the case. Then, by condition (3.10), this is true for all agents. But this contradicts condition (3.8) whenever \( \eta > 0 \) since \( \theta > 0 \) by (3.9).

Consider next the set of agents with non-binding enforcement constraints (\( \mu_i = 0 \)). For these agents, \( 1 + \lambda_i = \tau \) which implies that \( k_i = \frac{k}{1 + \lambda_i} \equiv (\frac{\alpha \tau}{\theta})^{\frac{1}{1-\alpha}} \). On the other hand, agents whose participation constraint is slack (\( \lambda_i = 0 \)) employ capital \( \frac{k}{1 + \mu_i} \equiv (\frac{\alpha}{\theta})^{\frac{1}{1-\alpha}} < k_i \), where the inequality follows from the fact that \( \tau > 1 \) by (3.10) since \( \mu_i + \lambda_i > 0 \) for almost all agents.
Finally, agents for which both constraints bind employ

\[ k_i = \left( \frac{\alpha(\tau - \mu_i)}{\theta} \right)^{\frac{1}{1-\alpha}} = (a_i^\alpha + \eta)^{\frac{1}{\alpha}} \in [\bar{k}, \bar{k}] \]

There only remains to show that these three groups of agents are separated by certain endowment thresholds. If agent \( i \in [0, 1] \) is in the group with non-binding enforcement constraints, \( \bar{k}^\alpha - a_i^\alpha < \eta \). Similarly, being in the group with non-binding participation constraints implies \( k^\alpha - a_i^\alpha > \eta \). Finally, a necessary condition for being in the group where both constraints bind is given by \( k_i^\alpha - a_i^\alpha = \eta \) for some \( k_i \in [\bar{k}, \bar{k}] \). These three conditions are mutually exclusive and define the thresholds we need.

Figure 3.2 shows the shape of the optimal capital allocation. Agents with low endowments have a binding enforcement constraint, but strictly prefer to participate in the optimal arrangement. They all employ the same level \( \bar{k} \) of capital. Conversely, agents with high endowments have a binding participation constraint but a loose enforcement constraint. These agents operate at the highest level of capital \( \bar{k} \). Agents in the middle have both constraints binding and operate with the capital stock such that their income level \( k_i^\alpha - \eta \) exactly matches their autarky income.

The optimal allocation can then be described by three equations that pin down the optimal level of enforcement \( \eta \) and the optimal allocation of capital by determining the two cut-off points \( \bar{a} \) and \( \bar{a} \),

\[
g'(\eta) = \int_{\{i|a_i \leq \bar{a}\}} \frac{1}{\alpha k_i^{\alpha-1}} di - \int_{\{i|a_i \leq \bar{a}\}} \frac{1}{\alpha k_i^{\alpha-1}} di \]  
\[ g(\eta) = \int a_i di - \int k_i di \]  
\[ \eta = \int_{\{i|a_i \geq \bar{a}\}} (a_i^\alpha - \bar{a}^\alpha) di. \]  

The first condition equates the marginal costs and the marginal benefits of enforcement expressed in units of capital input. More enforcement enables a better allocation of capital across agents. Hence, one can achieve the same output with a lower aggregate input of capital. Interestingly, marginal benefits can then be expressed as the wedge between the inverse of the marginal product of capital of unconstrained and constrained agents.\(^{11}\) The other two

\(^{11}\)This is reminiscent of an inverse Euler equation describing efficiency in the literature on Mirleesian taxation in dynamic economies. See e.g. Rogerson (1985) or Kocherlakota (2005).
equations describe the feasibility of allocating capital and of transfers in terms of the two endowment cut-offs.

The optimal allocation leads to a more equal income and consumption distribution than under autarky. In particular, all agents above the lower endowment threshold $a$ receive their autarkic income, while all other agents receive a fixed income higher than autarky. Hence, improving institutions by investing into costly enforcement benefits agents at the lower end of the endowment distribution.

**Proposition 3.4.** The optimal income distribution is a left-censored version of the income distribution under autarky. Specifically,

$$k_i^\alpha - t_i = \begin{cases} a^\alpha & \text{if } a_i \leq a \\ a_i^\alpha & \text{if } a_i \geq a. \end{cases}$$

**Proof.** Agent $i$’s end-of-period income is $k_i^\alpha - t_i$ for all $i \in [0,1]$. All agents whose assets exceed $a$ have a binding participation constraint. Hence, they have the same income as under autarky. Agents with endowments under $a$ all realize income $k^\alpha - \eta = a^\alpha$. □

It is instructive to compare the optimal allocation when enforcement is endogenous to the
allocation that prevails when enforcement is costless and exogenous. In both cases, positive enforcement reduces the dispersion of marginal products vis-a-vis autarky by making capital use constant past a certain asset threshold. There are, however, two crucial differences when the optimal allocation involves a costly enforcement choice. First, the planner also equates capital use among agents with low endowments. Second, transfers after production are optimally set such that only poor agents gain from enforcement, with all other agents receiving the same income as in autarky. This suggests that the planner combines investment in enforcement institutions with some redistribution of endowments when covering the costs of this investment.

3.3 Implementing the Optimal Allocation

To make these ideas precise, we will now explain how the optimal allocation can be implemented with decentralized capital markets. In order to implement the optimal allocation, it is necessary, first, to finance the introduction of enforcement at the desired level $\eta$. Denote agent $i$’s contribution to the aggregate enforcement cost by $\kappa_i$. One could think of this as an entry fee (or subsidy, when $\kappa_i < 0$) for participating in the market for capital. The planner invests these fees into establishing enforcement level $\eta = g \left( \int \kappa_i di \right)$. Then agents enter capital markets with a new endowment position equal to $\hat{a}_i = a_i - \kappa_i$ and trade capital at a competitively determined rate $R$ subject to a borrowing constraint given by

$$(k_i - \hat{a}_i)R \leq \eta.$$ (3.14)

The following result shows that the optimal allocation can be implemented via decentralized markets provided the cost of enforcement is distributed in a specific fashion.

**Proposition 3.5.** Let $(\eta, k, t)$ be the solution to the social planner’s problem. Define $R = \alpha k^{\alpha - 1}$ and let $\kappa_i = a_i - (k_i - \frac{t_i}{R})$ for all $i$. Given the contribution schedule $\{\kappa_i | i \in [0, 1]\}$, competitive markets implement the optimal solution with the equilibrium interest rate given by $R$.

**Proof.** Let $\hat{a}_i = a_i - \kappa_i$ for all $i$ so that the total supply of capital available for production is $\int \hat{a}_i di$. Then, the candidate allocation clears the capital market. Indeed,

$$\int \hat{a}_i di = \int k_i di - \frac{1}{R} \int t_i di = \int k_i di$$

14
since \( \int t_i \, di = 0 \) at the planner’s solution. Note that this also implies that the contribution schedule \( \kappa_i \) covers the enforcement cost \( g(\eta) \).

We only have to verify that agent \( i \) chooses the optimal capital input, \( k_i \), at interest rate \( R \). If agents are unconstrained, they will choose a capital level such that the marginal product is equal to \( R \). Constrained agents, however, will choose a capital level that satisfies the constraint (3.14).

Consider an agent with \( a_i > \bar{a} \) who, therefore, has \( k_i = \bar{k} \) and \( t_i < \eta \) in the optimal allocation. In that case, \( (\bar{k} - \hat{a}_i)R = t_i < \eta \) so that, given \( R \), the agent does choose to operate at \( k_i = \bar{k} \).

Agents with \( a_i \leq \bar{k} \) are constrained in the optimal allocation so that \( t_i = \eta \). By the definition of \( \kappa_i \), this implies that \( \hat{a}_i = k_i - \frac{\eta}{R} \). Hence, the agent chooses \( k_i < \bar{k} \) given his borrowing constraint (3.14), as needed. This completes the proof.

One could interpret \( \kappa_i \) as a tax schedule to finance the establishment of a capital market. This tax schedule, however, is no monotonic in initial endowments, as we now argue.\(^{12}\)

**Corollary 3.6.** Assume that \( \eta > 0 \). The contribution schedule \( \{\kappa_i| i \in [0, 1]\} \) rises monotonically with endowments on \( [0, \bar{a}] \), decreases on \( [\bar{a}, \bar{k}] \), and rises once again past \( \bar{k} \). Furthermore, the schedule’s local maximum at \( a \) is strictly positive, while its local minimum at \( \bar{k} \) is zero. In particular, \( \kappa_i > 0 \) whenever \( k_i < a_i \).

**Proof.** If \( i \) is such that \( a_i < \bar{a} \), then \( \kappa_i = a_i - (\bar{k} - \frac{\eta}{R}) \), an expression which rises with \( i \) since \( a_i \) does. If \( i \) is such that \( a_i \in [\bar{a}, \bar{a}] \), then

\[
\kappa_i = (a_i - k_i) + \frac{\eta}{R} = a_i - (a_i^\alpha + \eta)^\frac{1}{\alpha} + \frac{\eta}{R}.
\]

Differentiating this expression with respect to \( a_i \) shows that the schedule decreases with endowments in this region. Finally, if \( i \) is such that \( a_i > \bar{a} \), then \( \kappa_i = a_i - (\bar{k} - \frac{\eta}{R}) \). Since participation binds for these agents, \( t_i = \bar{k}^\alpha - a_i^\alpha \). Some algebra then shows that

\[
\kappa_i R = \bar{k}^\alpha - \bar{k} R - (a_i^\alpha - a_i R).
\]

\(^{12}\)We would like to thank Christian Zimmerman for pointing this out to us.
Because $\bar{k}$ is the value of $k$ that maximizes $k^\alpha - kR$, it follows that $\kappa_i$ falls with $a_i$ until $a_i = \bar{k}$, where it is zero, and then rises.

Since $\kappa_i = 0$ when $a_i = \bar{k}$ and the compensation schedule decreases on $[a, \bar{k}]$, it follows immediately that the schedule’s local maximum at $a$ is strictly positive.

This confirms that the planner achieves a better allocation of capital by using investment in institutions and redistribution in a complementary way. Some agents bear a disproportionate fraction of the institutional cost, a set which includes agents at the top of the distribution. The ensuing trade, however, enables high-endowment agents to exactly recuperate their contribution.

4 Inequality and Institutions

4.1 More Inequality in Marginal Products Leads to Better Institutions

Investments in institutions are only necessary when there is some measurable inequality in endowments. More generally, whenever endowments are more unequally distributed, the ex-ante dispersion in marginal products increases. Hence, one would expect that economies with more endowment inequality should invest more resources in institutions that provide better enforcement.

To formalize this, we model the notion of rising inequality as follows. Let $E(a)$ denote the average endowment. We say that the endowment schedule $\hat{a} = a + \delta(a - E(a))$ is more unequal than the distribution $a$ when $\delta > 0$. Throughout this section, we assume that for $\delta$ small enough, endowments remain almost surely non-negative. Alternatively, we could produce symmetric results for mean-preserving contractions rather than spreads.

We first look at the case where there are only two agents with different endowments. In that simple case, the impact of greater inequality on returns to enforcement is transparent. The nature of the optimal allocation is easy to describe, as the two cut-off points determine the allocation. When the spread in endowments increases, at the old enforcement level the spread in capital inputs must also increase. But then the marginal benefit of enforcement exceeds the marginal costs. Hence, it is optimal to invest more in enforcement.

Proposition 4.1. Assume that there are two agents. When the distribution of endowments
becomes more unequal, the optimal level of enforcement $\eta$ increases.

Proof. Write initial endowments in this case as $(a - \delta, a + \delta)$ where $a > 0$ and $\delta \in [0, a)$. Denote the production function as $f$ and its inverse as $h$. Since $f$ is strictly concave, $h$ is strictly convex. We also denote capital use by $k_1$ for the endowment-poor agent, while $k_2$ denotes capital use by the endowment-rich agent.

Suppose the enforcement level is given by $\eta > 0$. Then the capital allocation must solve

$$k_2 = h(f(a + \delta) - \eta)$$

where $k_2 \leq 2a$. Indeed, the enforcement constraint has to be binding for the poor agent at the optimal choice since otherwise enforcement costs could be reduced. This means that the rich agent receives a transfer of exactly $\eta$ at the optimal arrangement.

The agent with the low endowment then operates with capital $k_1 = 2a - h(f(a + \delta) - \eta) - g(\eta)$. Hence, total output is given by

$$\Phi(\eta, \delta) \equiv f(a + \delta) - \eta + f(2a - h(f(a + \delta) - \eta) - g(\eta)).$$

Note that the function $\Phi$ is strictly concave in $\eta$. The envelope theorem then implies that at the optimal level of enforcement

$$\frac{\partial \eta}{\partial \delta} = -\frac{\Phi_{12}(\eta(\delta), \delta)}{\Phi_{11}(\eta(\delta), \delta)}$$

Differentiating the function $\Phi$ with respect to $\eta$ we obtain

$$\Phi_1(\eta(\delta), \delta) = -1 + f'(k_1)(h'(f(a + \delta) - \eta) - g'(\eta))$$

where $k_1 = 2a - h(f(a + \delta) - \eta) - g(\eta)$ is the capital allocated to the low endowment agent.

We then have that

$$\Phi_{11}(\eta(\delta), \delta) = f''(k_1)(h'(\cdot) - g'(\cdot))^2 - f'(k_1)(h''(\cdot) + g''(\cdot)) < 0,$$

since $h$ is convex and $f$ is strictly increasing and concave. Furthermore, both $h(f(a + \delta) - \eta)$ and $h'(f(a + \delta) - \eta)$ are increasing in $\delta$, as $f(a + \delta)$ rises with $\delta$. This implies that $k_1$ decreases
in $\delta$ and the concavity of $f$ implies that $f'(k_1)$ rises with $\delta$ too. Thus we obtain that

$$\Phi_{12}(\eta(\delta), \delta) > 0$$

which completes the proof. \qed

Returning to the case with a continuum of agents, suppose that the planner chooses to bear the fixed cost $\lim_{\eta \downarrow 0} g(\eta)$ and to invest in strictly positive enforcement for a given distribution of endowments. We now show that this remains the case if the distribution of endowments becomes more unequal.

**Proposition 4.2.** Assume that the planner opts for strictly positive enforcement for a given endowment distribution. This remains true when the endowment distribution becomes more unequal.

**Proof.** Let $k$ be the optimal capital allocation in the first economy while $\eta > 0$ is the chosen degree of enforcement. We must have that $\int k_i^\alpha di \geq \int a_i^\alpha di$.

Consider now the more unequal distribution of endowments described by $\hat{a}_i = a_i + \delta(a_i - a^*)$. We will show that holding $\eta$ fixed a feasible capital allocation $\hat{k}$ exists in the more unequal economy such that $\int \hat{k}_i^\alpha di \geq \int \hat{a}_i^\alpha di$. Hence, strictly positive enforcement remains optimal for a more unequal distribution of endowments.

For all $i$, let

$$\hat{k}_i^\alpha = k_i^\alpha + \hat{a}_i^\alpha - a_i^\alpha$$

where it is assumed that $\delta$ is small enough that $\hat{k}_i \geq 0$ for all $i$. This is without loss of generality as the argument we use below is local. Leaving transfers unchanged, participation is clearly since it was in the original economy. The new allocation is better than autarky, since

$$\int \hat{k}_i^\alpha di = \int \hat{a}_i^\alpha di + \int k_i^\alpha di - \int a_i^\alpha di \geq \int \hat{a}_i^\alpha di.$$

We need to show that the new allocation satisfies the resource constraint. First, note that the total capital employed is given by

$$\int \hat{k}_i di = \int ((k_i^\alpha - a_i^\alpha) + \hat{a}_i^\alpha)^{1/\alpha} di.$$
Differentiating the integrand with respect to $\delta$ gives (up to multiplying constants)

\[
((k_i^\alpha - a_i^\alpha) + \dot{a}_i^\alpha)^{1/\alpha - 1} (a_i + \delta(a_i - a^*))^{\alpha - 1} (a_i - a^*).
\]

Hence, evaluating this expression at $\delta = 0$ shows that small changes to $\delta$ do not increase the total capital employed provided

\[
\int \left( \frac{k_i}{a_i} \right)^{1-\alpha} (a_i - a^*) di
\]

is weakly negative. We know that the original optimal allocation is such that $\left( \frac{k_i}{a_i} \right)^{1-\alpha}$ decreases as $i$ rises. Hence, by Chebyshev’s integral inequality, we have

\[
\int \left( \frac{k_i}{a_i} \right)^{1-\alpha} (a_i - a^*) di \leq \int \left( \frac{k_i}{a_i} \right)^{1-\alpha} di \times \int (a_i - a^*) di = 0
\]

which completes the proof.

Intuition suggests that monotonicity holds in this environment in a more general sense. Increases in the dispersion of marginal products should lead the planner to increase $\eta$. Looking at equations (3.11) through (3.13), notice that for any given enforcement level $\eta$, the last two equations alone pin down the cut-off levels for the optimal capital allocation. Hence, one can find this allocation for any value of $\eta$ independently of the enforcement choice. In other words, one can solve for the optimal level of enforcement by first determining the optimal capital allocation as a function of $\eta$ and, then, compare the marginal cost and benefits to find the efficient level of enforcement. Hence, it is optimal to raise the enforcement level for any marginal change in the distribution of endowments that causes more agents to be constrained or the wedge between the two cut-off points to increase.

### 4.2 The Engerman-Sokoloff Hypothesis

Our framework predicts that economies with more endowment inequality should be quicker to invest in institutions that support trade. This prediction may seem puzzling in light of the historical evidence for the Western Hemisphere surveyed by Engerman and Sokoloff (2002). Their evidence suggests that human and physical capital were more highly concentrated in Latin America than in the United States and Canada, and that the latter countries chose
to invest in institutions conducive to trade much earlier. In this section we argue that the efficiency considerations we emphasize in this paper, far from casting doubt on this hypothesis, provide additional support for it.

To see this, augment our static model to include heterogeneity in both physical and human capital. Agents are now endowed with a quantity $a_i > 0$ of physical capital and a level $h_i > 0$ of human capital for all $i \in [0, 1]$ with the joint distribution of human and physical capital described by $\mu$. Agents are also endowed with a technology that transfers physical capital into consumption goods according to a Cobb-Douglas production function $h^{1-\alpha}k^\alpha$, where $h$ is the human capital of the agent and $\alpha \in (0, 1)$. Since human capital cannot be traded across agents, the endowment of human capital acts like an agent-specific productivity parameter that is fixed.

To show that efficiency considerations can indeed rationalize the Engerman and Sokoloff hypothesis, we simply consider the case where the choice is between autarky and full enforcement. If there is no investment in enforcement, default cannot be punished hence all transfers are zero and autarky prevails. Aggregate output is then given by

$$y^A = \int h^{1-\alpha}a^\alpha d\mu = E(h^{1-\alpha}a^\alpha) = E(h^{1-\alpha})E(a^\alpha) + COV(h^{1-\alpha}, a^\alpha). \quad (4.1)$$

Alternatively, the planner can achieve full enforcement at an aggregate fixed cost of $C > 0$ units of capital. Markets are then complete and the planner is able to equate marginal products across agents. Denoting the total endowment of human and physical capital by $\bar{h}$ and $\bar{k}$ respectively, we obtain, for all $i \in [0, 1]:

$$\frac{k_i}{h_i} = \frac{\bar{k}}{\bar{h}}. \quad (4.2)$$

Since with full enforcement all transfers can be enforced, participation constraints for all agents can be met if and only if aggregate output increases after the enforcement cost $C$ has been incurred. Using aggregate resource feasibility, it follows directly that full enforcement leads to aggregate output equal to

$$y^E = \int h_i \left(\frac{\bar{k}}{\bar{h}}\right)^\alpha d\mu = \left(\int h d\mu\right)^{1-\alpha} \left(\int a d\mu - C\right)^\alpha = E(h)^{1-\alpha}(E(a) - C)^\alpha. \quad (4.3)$$

\textsuperscript{13}An optimal allocation with enforcement can once again be implemented via decentralized markets with a specific schedule of entry fees and subsidies, where the fee schedule has to satisfy all agents’ participation constraints.
The planner will choose to invest in enforcement whenever \( y^E > y^A \).

Holding the endowment distribution of the other factor fixed, a mean preserving spread in either human capital or physical capital endowments lowers output under autarky without affecting the complete market outcome. Hence, as in the analysis with only one input, more inequality can lead to more investment in institutions, as the benefits from trade have increased. However, in this two-dimensional setting, for any given marginal distribution of physical and human capital, the correlation in the endowments of both factors of production also determines institutional investment. If endowments in human and physical capital are sufficiently positively correlated, aggregate output is higher under autarky, and there is no institutional investment.

**Proposition 4.3.** Introducing complete markets with full enforcement at a fixed cost \( C \) leads to higher output than under autarky if and only if

\[
E(h^{1-\alpha})(E(a) - C)^\alpha - E(h^{1-\alpha})E(a^\alpha) > COV(h^{1-\alpha}, a^\alpha). \tag{4.4}
\]

The intuition for this result is straightforward. In economies where both human and physical capital are highly concentrated, the gains from introducing institutions are small, as the marginal products of capital are not very unequally distributed. In fact, when \( h \) and are \( a \) are perfectly correlated, we have

\[
COV(h^{1-\alpha}, a^\alpha) = E(h)^{1-\alpha}E(a)^\alpha - E(h^{1-\alpha})E(a^\alpha),
\]

so that inequality (4.4) can never be met.

This result underscores the fact that it is inequality in marginal products before trade that matters for returns to institutional investments, not endowment inequality per se. In the case of the Western Hemisphere, the fact that both human and physical capital have been highly concentrated historically in much of Latin America could explain why institutional quality has lagged behind its counterpart in the United States or Canada. In environments where physical and human capital endowments are highly correlated, institutions conducive to trading physical resources may not have much effect on output and growth, unless poor individuals are able to acquire more human capital.
5 The Dynamics of Enforcement and Inequality

How do institutions and inequality evolve over time given initial conditions? This section provides some answers by introducing simple inter-generational linkages in our model. Time is discrete and denoted by $t \in \{0, 1, \ldots\}$. In every period $t \geq 0$, a single member of each family $i \in [0, 1]$ is alive and inherits as endowment a given fraction $\gamma_i \in [0, 1]$ of their parent's income, where $\gamma_i$ is strictly increasing in $i$. This endowment can be used in production exactly as in the static model, or as investment in the enforcement technology. We also assume that investments in enforcement fully depreciate across periods, although this could be relaxed with little effect on our results. For simplicity, we will also assume that the current enforcement choice is myopic in the sense that each generation makes their institutional choice without taking into account the effects of the current choice on future generations’ welfare.\(^{14}\)

If there is no enforcement technology, the only feasible allocation of capital is autarky in all periods. It follows directly that in that case the endowment of members of lineage $i$ converges geometrically to $\gamma_i^{1-\alpha}$ over time. Hence, the endowment distribution converges to the corresponding invariant distribution.\(^{15}\) We assume that in period 0, the distribution of endowments is this invariant distribution, which amounts assuming that lineages have lived in autarky for long enough prior to that time to be near their invariant autarkic state.

With the possibility to invest in enforcement, we assume that this invariant distribution is sufficiently unequal that it is optimal to bear the fixed cost in period 0. Otherwise, no investment in enforcement is ever made, and the economy remains forever at the initial invariant distribution. Under this assumption, bearing the fixed cost remains optimal in all subsequent periods, and introducing enforcement leads to a progressive reduction of endowment inequality. This in turn reduces the benefits of investing in enforcement and successively reduces such investment.

\(^{14}\)Positing exogenously given intergenerational linkages is standard in most of the existing literature on inequality, growth and missing markets. See for example Aghion and Bolton (1996), Banerjee and Newman (1993) or Benabou (1996). There are at least two interpretations for such transfers: “warm-glow” altruism in the sense of Andreoni (1989), and intergenerational spillovers. The second interpretation is best understood if one thinks of productive resources in part as human capital. It has the advantage of side-stepping an obvious weakness of the warm-glow interpretation, namely the fact that lineages fail to internalize the consequences of their transfers on the welfare of their offspring. That concern is particularly strong in environments with redistribution policies. We thus emphasize the spillover interpretation and assume that generations are not linked in any other way.

\(^{15}\)Note that the endowment distribution converges at a geometric rate to a single mass point if lineages bequeath the same fraction of their income regardless of whether institutional investments are made. Eventually then, output is at its maximum independently of the economy’s history of institutional choices.
Proposition 5.1. If it is optimal to invest in enforcement at date \( t = 0 \), then the optimal allocation features a positive enforcement level for all periods \( t \geq 0 \) that decreases over time. Furthermore, the economy converges monotonically to a long-run invariant distribution of income and endowments with progressively less inequality and higher output.

Proof. Let \( a^t \) be the endowment function and \( \eta_t \) the optimal enforcement level in period \( t \). We proceed in two steps. First, we show that enforcement decreases in period 1, \( 0 < \eta_1 < \eta_0 \). Then, we establish that endowments increase over time, i.e. that \( a^2_i \geq a^1_i \) for all \( i \). The desired result will then follow by induction.

Given the initial endowment distribution \( a^0_i = \gamma_i^{1-\alpha} \), the optimal allocation at \( t = 0 \) is described by equations (3.11)-(3.13). There are two cut-off points \( \underline{a}^0 = \gamma_0^{1-\alpha} \) and \( \bar{a}^0 = \gamma_0^{1-\alpha} \), determining capital and transfers given the optimal level \( \eta_0 \). The new endowment function is then given for all \( i \in [0, 1] \) by

\[
a^1_i = \begin{cases} 
\gamma_i^{1-\alpha} & \text{if } \gamma_i \leq \gamma_0 \\
\gamma_i^{1-\alpha} & \text{if } \gamma_i \geq \gamma_0.
\end{cases}
\]

In particular, because \( \gamma_i^{1-\alpha} > \gamma_0^{1-\alpha} = a^0_i \) whenever \( \gamma_i < \gamma_0 \), we have that \( E(a^1) > E(a^0) \).

Suppose first that we constrain the planner to continue opting for enforcement level \( \eta_0 \) in period 1. It is straightforward to verify that the optimal allocation of capital is still described by equations (3.12) and (3.13). In particular, the new upper endowment threshold \( \bar{a}^1(\eta_0) \) is the same as in period 0, as the enforcement level has not changed. The lower threshold \( \underline{a}^1(\eta_0) \) has to increase from its period 0 value, as the aggregate endowment \( E(a_1) \) has gone up. Otherwise, some capital would not be used in production which cannot be optimal. These new thresholds pin down the new optimal allocation conditional on keeping the enforcement level at \( \eta_0 \).

We will first argue that this candidate allocation yields an average income level that exceeds its autarky counterpart at endowment distribution \( a^1 \), so that, \( \eta_1 > 0 \). First note that the optimal allocation of capital \( \{k^0_i : i \in [0, 1]\} \) in period 0 is still feasible in period 1 given enforcement level \( \eta_0 \). Indeed, for any \( i \in [0, 1] \) such that \( a^0_i \geq \gamma_0^{1-\alpha} \), \( a^1_i = a^0_i \) so the value of autarky is the same in period 0 and 1. For any \( i \) such that \( a^0_i < \gamma_0^{1-\alpha} \), we have

\[
(\underline{a}^0_i) - \eta = \gamma_0^{1-\alpha} > \gamma_i^{1-\alpha} = a^1_i(\gamma_i)^{\alpha},
\]

23
as \( \gamma_i < \gamma_0 \). This implies directly that in period 1 income is higher for everyone at the optimal period 0 allocation than it would be under autarky. As transfers sum to zero, aggregate output is also higher with enforcement than with autarky. Hence, it remains optimal to invest enforcement in period 1.

We will show next that \( \eta_1 < \eta_0 \). At \( \eta_0 \), we have for the new thresholds \( \underline{a}^1(\eta_0) \) and \( \overline{a}^1(\eta_0) \) that

\[
g'(\eta_0) > \frac{1}{\alpha} \int_{\{i | a_i \leq \overline{a}\}} \frac{1}{1 - \alpha} di - \frac{1}{\alpha} \int_{\{i | a_i < \overline{a}\}} \frac{1}{1 - \alpha} di,
\]

as the corresponding cut-off point for capital \( k_i(\eta_0) \) has increased. From the concavity of the objective function and the strict convexity of the constraint set – which is ensured by our assumptions on \( g \) conditional on incurring the fixed cost – we have a unique optimal value of enforcement \( \eta_1 \) that satisfies the first-order condition (3.11). Furthermore, for \( \eta \to 0 \) marginal benefits exceed marginal costs and for \( \eta \to \infty \) the opposite is true. Hence, optimal level of enforcement in period 1 must decrease, \( \eta_1 < \eta_0 \).

Finally, we will show that endowments increase over time for almost all \( i \). Note that the endowment distribution has not changed above the cut-off point \( \underline{a}_0 \). As \( \eta_1 < \eta_0 \), by equation (3.13) it must then be the case that \( \overline{a}^1(\eta_1) \) increases relative to its period 0 value. Suppose now that the lower cut-off point decreases, i.e. \( \underline{a}^1(\eta_1) < \underline{a}^0(\eta_0) \). Since all agents with \( a^1_i \geq \underline{a}^0(\eta_0) \) have the same endowment level as in period 0, their binding participation constraint implies that they have the same income level. All other agents have a strictly lower income level than \( \underline{a}^0(\eta_0) \). Since transfers sum to zero with the new threshold \( \overline{a}^1(\eta_1) \), it must be the case that total output has declined relative to its period 0 value. Since the old enforcement level \( \eta_0 \) is still feasible, this allocation cannot be optimal. A contradiction.

Hence, \( \underline{a}_1 > \underline{a}_0 \). A simple recursive argument then shows that there is positive, but declining enforcement in all subsequent periods. This implies that average income rises over time, rising the lower income threshold as well. We have then shown that the sequence of endowment distributions is a monotonically increasing sequence of distribution functions on \([0, 1]\). Furthermore, we can bound the values of each distribution of the sequence below by 0 and above by \( \gamma_{\text{max}} \). It then follows that the sequence of endowment distributions converges to some distribution as \( t \to \infty \). Along this sequence, we have successively less inequality and lower, but strictly positive enforcement.

This result implies that differences in institutional choices caused by differences in initial conditions can persist indefinitely, for reasons not unlike those explored in another context.
by Monnet and Quintin (2007). Institutional investments, in turn, allow for inequality to become reduced over time in a very specific sense. The endowment distribution is a censored version of the autarky distribution, with an ever higher censoring point. While the fixed cost must be borne ad infinitum, the fact that the ex-ante dispersion of marginal products falls over time make the optimal level of enforcement – or quality of institutions – fall over time as well.

6 Concluding remarks

In much of the recent literature on inequality and growth, it is the combination of endowment inequality and market imperfections that lead to bad economic outcomes and thereby create a rationale for redistribution. In this paper, we have pointed out that as long as it is possible to invest in better functioning markets, the optimal social arrangement calls for a combination of redistribution and institutional investment.

However, it is critical to recognize that it is ex-ante inequality in marginal products that matters, not endowment inequality per se. For instance, societies where physical and human capital endowments are highly correlated do not necessarily have an incentive to invest into costly institutions that improve market exchange. This suggests that investment in human capital and in institutional quality are intimately related. Investments in human capital raise the returns to investments in institutions. Conversely, better institutions enable nations to direct physical resources to their most productive uses, which raises returns to human capital development. Under those circumstances, the question of whether institution quality or human capital is the key to development success may be immaterial (see for example Glaeser, La Porta, Lopez-de-Silanes and Shleifer (2005) or Galor and Moav (2006)), as it is the dynamic interaction between the two that matters.

Overall, our work strongly suggests that in order to understand why some nations have persistently bad institutions, it is crucial to consider where the gains from these institutions arise, and how these gains can be redistributed throughout the economy.
7 Appendix

7.1 Existence and generic uniqueness

**Proposition 7.1.** A solution to the social planner’s problem exists. The solution is generically unique.

*Proof.* The planner’s problem is a Mayer problem with integral (isoperimetric) constraints (see section 4.8 in Cesari (1983)). In order to apply Filippov’s existence theorem, we need to restrict transfers and capital choices to a compact set. One can impose arbitrary bounds on both objects that are large enough not to bind at any solution. This ensures existence.

To establish generic uniqueness, note that given the fixed cost associated with implementing the enforcement technology, we need to compare the value of the problem when $\eta = 0$ – i.e., the value of the problem at autarky – and the value of the problem when the planner chooses to bear the fixed enforcement cost $\lim_{\eta \downarrow 0} g(\eta)$.

This second problem corresponds to solving the problem assuming that $g(0) = \lim_{\eta \downarrow 0} g(\eta)$. We will argue that the solution under that assumption is unique, so that the only case in which multiple solutions exist is when that solution happens to give exactly the same value as autarky. Generically therefore, there is at most one solution.

Under the assumption that $g(0) = \lim_{\eta \downarrow 0} g(\eta)$, the planner’s choice set is convex in $(\eta, k, t)$. Since the planner’s objective function is strictly concave in $k$, there is at most one optimal capital allocation in that case. The resource constraint then implies that $\eta$ must be unique as well. The transfer scheme is also unique because, as we argue below, either the participation or the enforcement constraint must bind for all agents. If the enforcement constraint binds, we have $t_i = \eta$. If the participation constraint of the agent is binding, transfers are given by $t_i = k_i^\alpha - a_i^\alpha$.

**References**


