Australian National University
Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Why Institutions Endure: Norms, Leadership, and the Difficulty of Reform

CAMA Working Paper 36a/2025
June 2025

Omer Majeed
Centre for Applied Macroeconomic Analysis, ANU

Abstract

How do institutions evolve? Why are they so persistent? And why are successful institutional transformations so rare, limited to outlier cases like Singapore, Türkiye, South Korea, Botswana, and China? This paper presents a new framework linking long-term institutional outcomes related to corruption and extractive practices to the dynamic interaction between population norms and leadership traits. The Population–Leadership Symmetry Principle posits that leadership traits reflect prevailing societal norms, as leaders emerge from within the population. Yet, meaningful institutional transformation requires a second mechanism: the Leadership Hysteresis Effect, where sustained, reformist leadership reshapes societal norms, embedding institutional change that persists beyond the leader's tenure. In both mechanisms, societal norms play a central role. For the Hysteresis Effect in particular, institutional reform depends on stability over time to gradually shift these underlying norms. In this framework, only long-duration and intensive leadership episodes generate durable improvements in governance; the model also explains why these reform episodes are rare. The model of this paper is calibrated to notable cases of institutional transformation. Empirically, I test the model using panel data and event studies, showing that societal corruption norms are strongly associated with leadership integrity over time. However, the absence of a valid external instrument limits causal inference; accordingly, the results are best interpreted as evidence of association rather than causation. Even so, the findings are robust across specifications and consistent with the model's predictions. Together, the findings offer a unified explanation for both institutional persistence and the conditions under which rare but lasting reform is possible.

Keywords

Institutions, societal norms, development, leadership, economic growth

JEL Classification

O43, D73, H83, P10, O57, Z18

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

Why Institutions Endure: Norms, Leadership, and the Difficulty of Reform *

Omer Majeed
Centre for Applied Macroeconomic Analysis, ANU

May 2025

Abstract

How do institutions evolve? Why are they so persistent? And why are successful institutional transformations so rare, limited to outlier cases like Singapore, Türkiye, South Korea, Botswana, and China?

This paper presents a new framework linking long-term institutional outcomes related to corruption and extractive practices to the dynamic interaction between population norms and leadership traits. The Population–Leadership Symmetry Principle posits that leadership traits reflect prevailing societal norms, as leaders emerge from within the population. Yet, meaningful institutional transformation requires a second mechanism: the Leadership Hysteresis Effect, where sustained, reformist leadership reshapes societal norms, embedding institutional change that persists beyond the leader's tenure.

In both mechanisms, societal norms play a central role. For the Hysteresis Effect in particular, institutional reform depends on stability over time to gradually shift these underlying norms. In this framework, only long-duration and intensive leadership episodes generate durable improvements in governance; the model also explains why these reform episodes are rare. The model of this paper is calibrated to notable cases of institutional transformation. Empirically, I test the model using panel data and event studies, showing that societal corruption norms are strongly associated with leadership integrity over time. However, the absence of a valid external instrument limits causal inference; accordingly, the results are best interpreted as evidence of association rather than causation. Even so, the findings are robust across specifications and consistent with the model's predictions. Together,

*I would like to thank Flavio Menezes, Guenther Schulze, Petr Sedlacek, Gabrielle Watson, John Asker, Gabriele Gratton, Jenny Gordon, Warwick McKibbin, Jane Golley, Beth Webster and participants at various seminars and conferences for their comments. This work has been influenced by the ideas of S. M. W. Ahmad, Eugene Wigner, Abdus Salam and M. A. Majeed. Disclaimer: The views expressed in this paper are those of the author and do not necessarily reflect the views of any affiliated institutions.
Contact information: omer.majeed@anu.edu.au

the findings offer a unified explanation for both institutional persistence and the conditions under which rare but lasting reform is possible.

Keywords: *Institutions, societal norms, development, leadership, economic growth.*

JEL Codes: *O43, D73, H83, P10, O57, Z18*

1 Introduction

How do institutions evolve, why are they so persistent, and what explains the rare transformations seen in outlier countries? Acemoglu et al. (2001) demonstrate that the persistence of extractive institutions—often marked by elite capture and corruption—is central to understanding the divergence between rich and poor economies. Corruption embedded within institutions systematically undermines development and long-term growth Mauro (2004); Hodge et al. (2011); Ashraf and Weil (2024); Kaufmann et al. (2010, 2005); Aidt (2009).

Acemoglu et al. (2001) emphasize the deep-rooted nature of institutions, particularly the long-lasting effects of colonial rule. While colonial rule may have had a significant impact on institutions, it may not fully explain cross-country variations. Historical examples suggest that institutions can be transformed under the right conditions. Countries such as Singapore, China, South Korea, Botswana, and Türkiye have experienced significant institutional shifts outside of colonialism. These examples highlight the need to better understand the conditions that enable such rare transformations.

While existing models (e.g., Acemoglu et al., 2001; North, 1990) emphasize institutional persistence, they do not explicitly model the dynamic interaction between leadership and population norms. This paper addresses this gap by introducing the *Population-Leadership Symmetry Principle* and the *Leadership Hysteresis Effect*, explaining both the mechanisms that reinforce institutional inertia and how sustained leadership can reshape societal values and entrench institutional change. While much of the literature emphasizes institutional persistence, this paper proposes a simple yet novel mechanism that not only accounts for such persistence but also explains how change can occur—and why it remains so rare.

Institutional inertia arises from the mirroring of leadership and societal values—a core feature of the Population–Leadership Symmetry Principle.[1] Accordingly, the model posits that enduring institutional change requires a shift in population norms.

[1]For the purposes of this paper, 'social' and 'societal' values are used interchangeably to refer to prevailing norms within a population.

Since these norms exhibit strong inertia, such transformation is rare. The Leadership Hysteresis Effect explains how a few countries—Singapore, China, South Korea, Botswana, and Türkiye—have achieved lasting reform by reshaping societal norms through sustained leadership. Such cases remain unaccounted for by Acemoglu et al. (2001), whose framework does not accommodate changes outside of colonial regimes.

Thus, while the model affirms that institutions are highly persistent, it also shows they are not static—aligning with North (1990) and Acemoglu and Jackson (2015), but contrasting with Acemoglu et al. (2001). This paper also aligns with Acemoglu and Robinson (2025) in suggesting that norms are an important part of institutional equilibrium. Through this approach, the paper not only deepens our understanding of how social norms shape institutional equilibria, but also provides a novel framework for examining how leadership traits and social norms influence institutional outcomes and policy effectiveness—filling an important gap in the literature (Acemoglu and Robinson, 2025).[2]

Paper in Brief and Related Background

I propose a simple framework to conceptualize how a country's institutions evolve through the interaction between population characteristics and leadership traits. Over time, societal and leadership characteristics form a self-reinforcing equilibrium, in which leadership traits related to corruption mirror societal norms—as leaders emerge from within the population and the society that shapes them. I term this dynamic the *Population–Leadership Symmetry Principle*. At the core of the symmetry principle is the idea that societal norms shape incentive structures, contagion effects, and habit formation—all of which influence individuals, including those who rise to leadership. Because institutions are ultimately composed of individuals, the prevailing societal norms strongly influence their performance.

The hypothesis that societal norms on corruption shape the attitudes and behaviors of the people in the society is bolstered by micro-level evidence. For instance, Gulino and Masera (2023) show that news of local corruption scandals increase the probability of theft at supermarkets in a significant manner, showcasing contagion effects of societal norms. Similarly, Ajzenman (2021) show that revelations of corruption by local officials lead to increased cheating by students of secondary schools on cognitive tests. There is also evidence that people in positions of power emulate the characteristics of their country. Fisman and Miguel (2007) show that diplomats from more corrupt countries accumulated significantly more unpaid parking violations in New York City. When diplomatic immunity from parking fines was removed, violations declined sharply—highlighting the roles of both habit formation and incentive respon-

[2]As Acemoglu and Robinson (2025, p. 682) note, 'much more is needed in the modeling of the joint evolution of culture, politics and institutions.'

siveness in corrupt behavior. Gächter and Schulz (2016) show through cross-sectional evidence from 23 countries that the prevalence of rule violations in society strongly determines individual honesty. At the suburb level, Keizer et al. (2008) provide experimental evidence that visible violations of social norms (e.g., graffiti, littering) encourage further norm violations. Bicchieri et al. (2022) show that societal values are an important determinant of norm compliance. Borjas (2000) finds that cultural norms, including attitudes toward corruption, persist across migrant generations.

This paper builds on this micro-level evidence by showing that, at the macro level, societal norms can significantly shape the traits of those who rise to leadership—highlighting a broader pattern of normative alignment between populations and their leaders.

Four key mechanisms form the basis of the symmetry principle.

First, as outlined in the principal-agent (PA) framework of this paper, prevalent corruption weakens enforcement and incentivizes corruption. High-corruption environments overwhelm enforcement systems, diluting enforcement capacity and lowering the probability that any individual case will be punished—thereby reducing the overall likelihood of effective enforcement. Strategic tolerance also emerges, including tit-for-tat protection among corrupt officials (Andvig and Moene, 1990).

Second, the social context heavily influences personal integrity and behavioral norms through habit formation and contagion effects. Individuals exposed to widespread rule violations are more likely to normalize and internalise corrupt behavior (Gulino and Masera, 2023; Ajzenman, 2021; Gächter and Schulz, 2016; Fisman and Miguel, 2007). In corrupt societies, corruption is thus fostered.

Third, the prevalence of integrity within the general population influences the selection of leadership. In societies with a larger share of non-corrupt individuals, it is more likely that future leaders, policymakers, judges, and law enforcement officials will be drawn from honest segments of society—a process referred to here as the selection pool effect.

Fourth, societal preferences exert normative pressure on leadership. In populations where values strongly favor honesty and good governance, there is greater demand for integrity in leadership and stronger constraints on corrupt behavior. Over time, when societal preferences and leadership behavior become aligned, a stable equilibrium emerges.

Together, these mechanisms lead to a substantive new conclusion: leadership traits mirror those of the population. The symmetry principle implies that in corrupt countries, it is not only that the leaders are corrupt but that corruption permeates the society. The symmetry principle thus reveals a dynamic of institutional persistence: corruption reinforces corruption. It explains how societal values shape the ethical profile of leaders—and why meaningful institutional change is rare without a shift in underlying norms.

Given this, how do leaders in rare cases implement successful reforms? In this framework, leadership emergence is stochastic over the long run. In corrupt societies, the probability of corrupt leadership is systematically higher. When reformist leaders emerge despite prevailing norms, meaningful institutional change can only take hold if leadership is sustained long enough to reshape the underlying societal values. This gives rise to the *Leadership Hysteresis Effect*—a process through which persistent and strong leadership gradually shifts population norms, embedding reforms that endure beyond a leader's tenure.

This effect underscores two key propositions: (1) societal transformation requires stability—only sustained and consistent leadership can entrench norm change and institutional reform; and (2) for institutions to improve in a durable way, the characteristics of the population must evolve—from corrupt to non-corrupt norms.

This framework aligns with North (1990), who argues that institutions evolve slowly and that informal institutions—people's beliefs and habits—often matter more than formal changes like constitutions or legal reforms. While the concept of leadership hysteresis is underexplored, Acemoglu and Jackson (2015) provide the closest view to this paper, showing how rare, prominent leaders can "counteract history" and shift social expectations across generations by coordinating beliefs, as in the case of Nelson Mandela. Relatedly, Giuliano and Nunn (2021) show that the long-term impact of institutional exposure—such as medieval self-governance—depends on its duration and intensity. This finding resonates with the Hysteresis Effect introduced in this paper, where both the duration and intensity of reformist leadership are essential to embedding institutional change. Further, evidence from successful cases of reform—such as Singapore, Türkiye, Botswana, China, and South Korea—is consistent with the Leadership Hysteresis Effect. In each case, reforms were sustained and intense enough to create lasting change that persisted beyond the leadership's tenure.

In the framework developed here, I use a difference equation setting to show how leadership can gradually reshape population characteristics. A central mechanism operates through exit rates and entry rates of corrupt and non-corrupt people. Exit rates reflect both natural attrition and institutionally driven removals. For example, a reformist leader can strengthen judicial integrity and independence, increasing the likelihood that corrupt individuals are identified and imprisoned—thereby accelerating the exit of corrupt agents from the system. This enforcement also serves as a public signal, shifting the incentives and behavior of potential entrants into the bureaucracy and political class. Over time, leadership can alter the entry rates of non-corrupt people by influencing education systems, signaling acceptable behaviors (including through media discourse), and value formation (linked through contagion and habit formation)—ultimately increasing the inflow of non-corrupt individuals.

The model is calibrated to reforms in Singapore, Türkiye, Botswana, and South Korea, capturing governance transitions and norm shifts. These calibrations support the central insight: that shifting societal norms is a gradual process requiring sustained, high-intensity leadership.

In summary, the framework argues that institutional persistence is the norm, as leadership traits tend to mirror prevailing societal values—a dynamic captured by the Population–Leadership Symmetry Principle. In contrast, the rare but durable reform episodes are explained by the Leadership Hysteresis Effect, whereby sustained, high-intensity leadership gradually reshapes societal norms. In both concepts, societal norms play a central role, serving as the unifying force that links the persistence of institutions with the rare pathways to their transformation. Crucially, societal norms not only shape leadership traits but also serve as the foundational channel through which institutional equilibrium is reached. The paper also formalises why the probability of such transformations is exceptionally low.

This paper complements Acemoglu and Robinson (2025), who argue that social norms are an important part of the equilibrium. While Acemoglu and Robinson (2025) conceptualize social change through a set of attributes—some evolving rapidly, others more slowly—-they describe their own framework as a 'sketch' and a 'preliminary step' toward understanding institutional change, and call for formal modeling. They suggest there needs to be a way to show how attributes are part of the equilibrium. This paper offers a novel framework to show that norms define equilibrium through a more formal model—grounded in the Population–Leadership Symmetry Principle.[3]

I start by using a simple PA framework to lay the micro foundations and showcase incentives at play in the society. PA models often analyze corruption through incentives like detection probability and rewards (Jain, 1998; Burguet et al., 2016), with some highlighting strategic complementarity, where corruption reinforces itself (Andvig and Moene, 1990). Existing models struggle to explain persistence of corruption (Mauro, 2004; Burguet et al., 2016) and overlook its societal impact when widespread (Persson et al., 2013). They also assume a fixed detection probability, treating institutions as static.

I extend Mauro (2004)'s two-period model to show how widespread corruption fosters inertia. Unlike standard PA models, I show how P_D and moral costs (m_i) can change based on the model parameters, linking it to the interplay between population and leadership traits. This partly addresses the limitations of PA models in contexts of widespread corruption (Persson et al., 2013), showing how incentives shift with the economy's corruption levels. The results of my PA model align with Becker (1968), which shows that incentives matter. Additionally, it aligns with micro-level studies

[3]The model examines norms across different segments of the population—i.e., societal norms and leadership norms. Using contagion effects, habit formation, incentives, and behavioral traits, the paper shows how norms are part of the institutional equilibrium. Using a difference equation framework, it demonstrates that corruption-related norms are slow-moving, in line with the perspectives of Nunn (2023) and North (1990). The model is also applied to, and calibrated using, historical case studies to examine notable institutional transitions across countries.

(Gulino and Masera, 2023; Ajzenman, 2021; Fisman and Miguel, 2007) suggesting that people—and therefore leaders—take their social cues from the broader society.

Summary of Empirical Results

There is a clear gap in the literature showing how norms can lead to institutional outcomes Acemoglu and Robinson (2025) and how policy or leadership can change them. This paper tries to fill this gap. The empirical analysis provides the first evidence in support of both the Population–Leadership Symmetry Principle and the Leadership Hysteresis Effect.

The empirical analysis for the Symmetry Principle draws on four strategies. I use person-level data from the World Values Survey (WVS) to construct country-level indicators of societal integrity, including attitudes toward tax evasion, fare-dodging, and benefit fraud. To study macro-level leadership characteristics on corruption, I use the Worldwide Governance Indicators (WGI) *control of corruption* index as the main proxy. This measure reflects the extent to which public power is used for private gain—capturing both petty and grand corruption, as well as elite capture—and aligns closely with the ethical dimension of leadership at the heart of the symmetry principle. As a robustness check, I also use the WGI's *Rule of Law* indicator, which captures some relevant leadership traits but is broader in scope, making it less central to the hypothesis.

The first estimation technique uses Ordinary Least Squares (OLS), which shows that weaker societal norms in earlier periods are associated with higher corruption in leadership in later periods. These results remain robust after conditioning on education, income per capita, trade openness, legal origins, democracy, freedom of expression, and regional governance trends. These variables may be linked with societal values and may directly impact leadership characteristics.

The second strategy further addresses reverse causality by using lagged societal values—approximately 10 to 20 years prior—to examine associations with future leadership outcomes using OLS.

The third strategy uses an alternative estimation technique, the system-GMM approach, to address endogeneity, reverse causality, country fixed effects, and unobserved heterogeneity. The findings remain robust.

Finally, the fourth strategy tests the robustness of the core findings using additional measures of corruption from the V-Dem dataset, including indicators based on public sector bribes and theft. These variables offer a distinct proxy for leadership corruption.

Across all four strategies, the results consistently indicate that societies with weaker honesty norms are more likely to experience corruption at the leadership level.

The main empirical limitation of this paper is the absence of a valid external instrument, which constrains the ability to draw causal inferences. As emphasized in Acemoglu and Robinson (2025) establishing causality in this domain is inherently challenging. Accordingly, the results should be interpreted as evidence of associations rather than causation. A promising direction for future research could be to identify natural instruments—possibly at the level of countries—which may enable causal identification, especially given the difficulty of finding suitable cross-country instruments.

Although the empirical focus is on the symmetry principle—which accounts for the broad alignment between societal norms and leadership traits—the analysis also provides the first cross-country evidence for outlier cases consistent with the *Leadership Hysteresis Effect*. This is done by leveraging quasi-experimental variation in leadership transitions across countries and contrasting sustained reformist leadership in Singapore (Lee Kuan Yew) and Türkiye (Atatürk) with short-lived episodes (e.g., Musharraf in Pakistan), using an event-study framework to trace lasting effects. The findings indicate that only long-duration, high-intensity leadership episodes are associated with persistent improvements in corruption and societal norms, consistent with the dynamics of the Leadership Hysteresis Effect. The section on the probability of change formalises why such cases remain exceptional.

Contributions

This paper develops a new framework for understanding how social norms and leadership traits jointly shape institutional equilibria—and how these elements can evolve over time. It contributes to multiple strands of literature, including institutional change, leadership dynamics, behavioral contagion, and the PA framework.

First, it proposes a dynamic framework in which a country's institutions evolve through the interaction of population characteristics and leadership traits, formalized through the *Population–Leadership Symmetry Principle*. As such the paper enriches the understanding between the interplay of social norms and leadership and how they can impact institutions.

Second, it introduces the *Leadership Hysteresis Effect*, capturing how sustained, reform-minded leadership can gradually shift societal norms and entrench institutional reform. The paper shows that even with discontinuous changes in leadership, norms on corruption evolve more gradually.

Third, the paper provides the first cross-country empirical evidence supporting both mechanisms. It connects recent micro-level findings in behavioral economics on contagious dishonesty and norm transmission—drawing on Gulino and Masera (2023), Fisman and Miguel (2007), and Bicchieri et al. (2022)—to show that these dynamics also operate at the national level.

Finally, the paper shows how the PA framework can be adapted to changing norms and different levels of corruption in the society.

The rest of the paper is structured as follows. Section 2 outlines the model setup. Section 3 develops the PA framework. Section 4 presents the main model. Section 5 calibrates the model to reform episodes. Section 6 presents data and results. Section 7 concludes.

2 Model set up and definitions

Corruption

Following Gulino and Masera (2023), I define corruption as the abuse of public power for private gain. I extend this definition to include individuals in the private sector who exploit positions of authority for unlawful personal benefit.

Population characteristics θ_t^{NC}

I assume that a country's population consists of both corrupt and non-corrupt segments, where θ_t^{NC} represents the proportion of the population that is not corrupt, while θ_t^{C} denotes its corrupt counterpart. As such, θ_t^{NC} can be understood as the first moment of the distribution of societal integrity—capturing the average propensity of the population to act honestly. Slightly less important though an alternate way to interpret θ_t^{NC}, is as an indicator of society's tolerance or aversion to corruption. For instance, a θ_t^{NC}value of 0.9 can be interpreted as a population that acts honestly 90% of the time, or, conversely, tolerates a small amount of corruption in the economy. In this framework distribution of norms is also important, though less explored in this paper.

Leadership Characteristics $C_{A,t}$

$C_{A,t}$ represents leadership characteristics not just by the top leadership, but other tiers of leadership including judges, policy makers, politicians etc[4] . I assume $C_{A,t}$ ranges from a lower bound $\underline{A} > 0$ to an upper bound $\overline{A} \leq 1$.

$$C_{A,t} \in (\underline{A}, \overline{A}] \tag{1}$$

A high level of leadership characteristics ($C_A = \overline{A}$) signifies a strong leadership commitment to good governance, anti-corruption measures and the development of

[4]With top leadership having more weight.

growth-enhancing institutions, while $C_A = \underline{A}$ represents the opposite—corruption and institutional decay. In part, $C_{A,t}$ is seen by public as a signal in what is acceptable in the society. So for instance, $C_{A,t} = 0.7$ can be interpreted as a strong, though not perfect, commitment to honesty—implying that approximately 70 percent of the leadership is non-corrupt. Extreme values of close to 0 or 1, are likely to be rare in this model.

Institutions

I consider an economy where institutions—characterized by the absence of corruption (denoted by $\widetilde{I}_t$) —are shaped by both leadership characteristics ($C_{A,t}$), and population characteristics (θ_t^{NC}).

Institutions can be conceptualized as complimentary inputs of two main factors (Fortunato and Panizza, 2015, Krieger, 2022). I assume a Cobb-Douglas [5] production function for institutions, given by:

$$\widetilde{I}_t = (\Psi C_{A,t}^{\lambda_t})\theta_t^{1-\lambda_t^{NC}} \tag{2}$$

where $0 \leq \lambda_t \leq 1$

Ψ is a leadership augmenting factor that determines how effectively leadership drives institutional change. Ψ could reflect leadership effectiveness, including internal harmony among layers of leadership and external factors [6] [7] .

3 Micro-foundations — Principal-Agent Model

This section provides a micro-foundation for how societal corruption incentivizes corrupt leadership, while honest societies reinforce integrity—supporting the Population–Leadership Symmetry Principle. In corrupt environments, weak detection and contagion effects enable systemic corruption, whereas in honest societies, institutional norms sustain accountability.

While the model is initially framed around agents in leadership roles—such as politicians, senior officials, or bureaucrats—the underlying incentive structure applies equally to the broader population. In this model, I assume that both leadership and citizens operate within the same institutional environment, responding to the same enforcement signals (P_D) and internalized moral costs (m_i).

[5] Institutions production function can also be made linear similar to KRIEGER, T. 2022. Democracy and the quality of economic institutions: theory and evidence. Public Choice, 192, 357-376. without loss of generality.

[6] For instance, under Lee Kuan Yew, this model would have high values for Ψ and C_A.

[7] Including geopolitical factors, access to markets and sanctions.

To explore the incentive dynamics, I apply the PA framework, highlighting how leadership, judicial enforcement (J) [8], contagion effects, and corruption incentives interact to shape governance outcomes. PA models often analyze corruption through incentives like detection probability and rewards (Burguet et al., 2016), with some highlighting strategic complementarity, where corruption reinforces itself (Andvig and Moene, 1990). Existing models struggle to explain corruption's persistence (Mauro, 2004; Burguet et al., 2016) and overlook its societal impact when widespread (Persson et al., 2013). They also assume a fixed detection probability, treating institutions as static.

PA Model

I extend Mauro (2004)'s two-period model to show how widespread corruption fosters inertia.

The model includes four stakeholders: the broader population (principal), agents, top leadership $C_{A,t}^M$, broader leadership $C_{A,t}$, and judiciary J, which enforces detection and influences both P_D and m_i through signaling.

Model Setup

- Principal: Citizens
- Agents: $C_{A,t}^M$, $C_{A,t}$
- Endowment: e received each period, independent of behavior
- Wages: w in each period if honest; if corrupt and caught in period 1, loses w in period 2
- Corruption payoff: Bribe b in each period if corrupt
- Moral cost of corruption: m_i per period; a bribe taken in the first period induces guilt in both periods
- Probability of being caught: $P_D = f(C_{A,t}, \theta_t^{NC})$, explained further below
- Penalty if caught: Must return all bribes received and pay an additional fine $\widetilde{T}$

[8] There are several ways to model J. $J_{t+1} = aJ_t + \beta C_{A,t}^M - \gamma\theta_t^C$. This says that there is judicial inertia, which can be overcome by strong leadership effort, but a corrupt population will weaken judicial reforms. However, for the purposes of this paper, I do not need to formalise J.

Utility Function

The agent's utility over the two periods is given by:

$$\text{Utility} = \frac{c_1^{1-\sigma} - 1}{1-\sigma} + \frac{c_2^{1-\sigma} - 1}{1-\sigma} \tag{3}$$

where c_1 and c_2 are consumption in periods 1 and 2, and σ is the inverse of the intertemporal elasticity of substitution. All income in each period is consumed.

If Honest

The leader's consumption is:

- Periods 1 and 2: $c_1 = c_2 = w + e$

Expected utility of an honest leader is:

$$V^{NC} = \frac{(w+e)^{1-\sigma} - 1}{1-\sigma} + \frac{(w+e)^{1-\sigma} - 1}{1-\sigma} \tag{4}$$

If Corrupt

- Period 1: $c_1 = w + e + b - m_i$
- Period 2:
  - with probability $(1 - P_D)$: $c_2 = w + e + b - m_i$
  - with probability P_D: $c_2 = e - b - \widetilde{T} - m_i$

The expected utility of a corrupt agent is:

$$\begin{aligned} V^C = &\frac{(w+e+b-m_i)^{1-\sigma} - 1}{1-\sigma} \\ &+ (1-P_D) \cdot \frac{(w+e+b-m_i)^{1-\sigma} - 1}{1-\sigma} \\ &+ P_D \cdot \frac{(e-b-\widetilde{T}-m_i)^{1-\sigma} - 1}{1-\sigma} \end{aligned}$$

Incentive Compatibility Constraint (IC)

The agents will choose to be honest if:

$$V^{NC} > V^{C}$$

Full condition:

$$\frac{(w+e)^{1-\sigma}-1}{1-\sigma} + \frac{(w+e)^{1-\sigma}-1}{1-\sigma} \geq \frac{(w+e+b-m_i)^{1-\sigma}-1}{1-\sigma} + (1-P_D)\cdot\frac{(w+e+b-m_i)^{1-\sigma}-1}{1-\sigma} + P_D\cdot\frac{(e-b-\widetilde{T}-m_i)^{1-\sigma}-1}{1-\sigma}$$

Simplified:

$$2(w+e)^{1-\sigma} > (1-P_D)(w+e+b-m_i)^{1-\sigma} + P_D(e-b-\widetilde{T}-m_i)^{1-\sigma}$$

Comparative Statics

- **Higher Bribe (b):** Increases corruption temptation.
- **Higher Probability of Detection (P_D):** Raises expected punishment, discouraging corruption.
- **Higher Moral Cost (m_i):** Raises internal guilt, reducing the net benefit of corruption.
- **Higher Penalty ($\widetilde{T}$):** Strengthens punishment, further deterring corruption.
- **Higher Wage (w):** Increases stable income, making honesty more appealing.

Incentives and Moral Costs as Corruption Varies

Probability of punishment can vary in a society based on corruption for two reasons. First, in corrupt societies, the high volume of corruption cases overwhelms the judiciary, stretching limited enforcement resources and reducing the likelihood that any individual case is detected and effectively prosecuted. This, in turn, lowers the perceived risk of punishment, thereby increasing the incentive to engage in corruption.

Second, since the judiciary is drawn from the population, it typically reflects prevailing norms θ_t^{NC}. Where widespread corruption leads to strategic tolerance, as individuals—including judicial officials—are themselves likely to be corrupt in corrupt societies. This creates "tit-for-tat" dynamics, where corrupt actors are reluctant to report others for fear of retaliation or mutual exposure (Andvig and Moene, 1990). Instead, these corrupt officials are more likely to extract side payments less than $\widetilde{T}$, diluting deterrence to corruption. As a result, enforcement becomes lenient or transactional, with side payments substituting formal penalties—further weakening deterrence and reinforcing systemic corruption. Additionally, corrupt officials are often likely to collaborate to extract resources together (Weisel and Shalvi, 2015). All of these factors lower the probability of P_D and incentivize corruption in corrupt societies.

In addition, moral costs (m_i) can vary in a society based on corruption for two reasons. First, due to habit formation and contagion effects in a corrupt society (Gulino and Masera, 2023; Ajzenman, 2021; Fisman and Miguel, 2007; Bicchieri et al., 2022), can breed further corruption as it lowers the moral cost of corruption for individuals.

In addition, pervasive corruption fosters a strategic mindset in which individuals rationalize that abstaining from illicit gains merely allows others to seize them—weakening internal moral constraints—much like African warlords who set up roadblocks, citing neighbouring warlords doing the same (Reno, 1998).

P_D and m_i and Endogeneity with Social Norms and Leadership Characteristics

Unlike standard principal-agent models, based on the above discussions I argue that P_D and m_i changes based on the population and leadership characteristics. This partly addresses the limitations of PA models in contexts of widespread corruption (Persson et al., 2013), showing how incentives and personal integrity (moral costs of corruption) shift with the economy's corruption levels. Specifically:

$$P_D, m_i = f(C_{A,t}^M, \theta_t^{NC}) \tag{5}$$

In normal times, when reformist leadership is absent, θ_t^{NC} influence both P_D and m_i. When $C_{A,t}^M$ and θ_t^{NC} are low, and then low detection rates further embolden corrupt behavior, worsening institutional inertia. The reverse is true for strong leadership and honest countries.

However, under strong reformist leadership—where $C_{A,t}^M > \theta_t^{NC}$, I assume the leader proactively reshapes the judicial system by appointing officials aligned with the reform agenda. This alters the composition of the judiciary j, increases the probability of detection P_D, and shifts the moral cost of corruption m_i through strong signaling.

Historical examples such as Atatürk in Türkiye, Lee Kuan Yew in Singapore exemplify this mechanism.

Equilibria in the Two-Period Principal-Agent Model with Probabilistic Enforcement

As such, I assume the probability of detection can take different values:

$P_D = P_l$ (low detection probability) — in a corrupt society, with weak governance, $P_l = \theta_t^{NC}$, when θ_t^{NC} is low. Similarly, m_i is lower when θ_t^{NC} is low.

Therefore, in corrupt societies, corruption is reinforced.

$P_D = P_h$ (high detection probability) — either society has strong anti-corruption characteristics, so θ_t^{NC} is high, or reformist (high $C_{A,t}^M$) alters J through strong and reformist appointments and therefore increases P_D. Similarly, high θ_t^{NC} will increase m_i.

Corrupt Equilibrium

P_l and low m_i — we end up with a corruption equilibrium, where $C_{A,t}$ is incentivized to act corruptly

$$2(w+e)^{1-\sigma} < (1-p_l)(w+e+b-c_i)^{1-\sigma} + p_l(e-b-\widetilde{T}-c_i)^{1-\sigma}$$

Honest Equilibrium

P_h and high m_i — we end up with a lowcorruption equilibrium, where $C_{A,t}$ is incentivized to be behave honestly

$$2(w+e)^{1-\sigma} > (1-p_h)(w+e+b-c_i)^{1-\sigma} + p_h(e-b-\widetilde{T}-c_i)^{1-\sigma}$$

Policy Interpretation

Corruption generates strategic complementarities that perpetuate corrupt behavior by distorting incentives and lowering moral costs. Conversely, when non-corrupt individuals are more prevalent, honesty is reinforced. These self-reinforcing mechanisms underpin the symmetry principle, whereby societal norms and leadership traits mirror one another over time.

4 Framework

Leadership–Population Symmetry Principle and Steady State

I argue that there exists symmetry between leadership and population characteristics, on the basis of four points.

First, as outlined in the PA framework, corruption is incentivized in societies where it is widespread. In such settings, systemic corruption overwhelms the judiciary, weakening enforcement and reducing the probability that any individual case will be prosecuted. Tit-for-tat tolerance among corrupt actors further undermines deterrence, as mutual complicity discourages reporting and punishment. Together, these dynamics erode accountability and reinforce the incentives for corruption.

Second, social context shapes individual integrity, habit formation, and contagion effects—both among the general population and within leadership. Exposure to corruption lowers the moral cost of corrupt acts, as individuals come to normalize such behavior, and where greater exposure to rule violations is associated with reduced personal integrity (Gulino and Masera, 2023; Ajzenman, 2021; Fisman and Miguel, 2007; Gächter and Schulz, 2016). As corruption becomes more pervasive, individuals are increasingly likely to internalise these behaviors, reinforcing their acceptance through social contagion and norm adaptation—sometimes even rationalising corrupt behavior as a defensive response to avoid being left behind.

Third, θ_t^{NC} captures the share of the population that adheres to non-corrupt norms, forming the pool from which leaders, policymakers and judges are drawn. A higher θ_t^{NC} increases the probability that leadership will be from the non-corrupt segment of the population — this I label as the selection pool effect.

Fourth, if θ_t^{NC} is high, then the population has higher preferences for honest, accountable and effective governance, therefore the population will exert pressures on $C_{A,t}$ to be non-corrupt. As such, $\theta_t^{NC} \approx C_{A,t}$ represents harmony between the population's preferences and that of its leaders, and therefore a steady state in the long run.

Taken together, these findings underscore that social norms function as powerful behavioral anchors. In sum, the symmetry principle posits that the characteristics of leadership tend to mirror the prevailing norms of the population, shaped through incentives, social contagion, the selection pool effect, and normative preferences. This dynamic gives rise to a self-reinforcing institutional equilibrium.

Societal Norms and Leadership Alignment

The symmetry principle as such argues that societies with higher θ_t^{NC} are more likely to produce leaders committed to honest behaviors, leading to the following relationship:

$$P(\text{Leadership is non-corrupt}) = E(\theta_t^{NC})$$

And since leadership characteristics are represented by $C_{A,t}$, we have:

$$P(C_{A,t}) = \theta_t^{NC} \tag{6}$$

This expresses the core intuition behind the Population–Leadership Symmetry Principle: over time, the average characteristics of the leadership will tend to mirror the average societal norms. This implies that the long-run population characteristics and norms, denoted by denoted by $\overline{\overline{\theta_t^{NC}}}$—representing the long-run average of θ_t^{NC} over a substantial historical period—ultimately shapes the long-run average of characteristics of the leadership, $\overline{\overline{C_{A,t}}}$.

The intuition behind the symmetry principle is that in societies with prevailing corrupt norms—or a high share of corrupt individuals—future leaders are more likely to be corrupt as well.

$$\overline{\overline{C_{A,t}}} \approx \overline{\overline{\theta_t^{NC}}} \tag{7}$$

Further, in the steady state we have:

$$\theta_t^{NC} = \overline{\overline{\theta_t^{NC}}} \approx \overline{\overline{C_{A,t}}} = C_{A,t} \tag{8}$$

Figure 1 illustrates this graphically. The x-axis represents the proportion of the non-corrupt population, while the curve shows the distribution around the mean (θ_t^{NC}). The y-axis gives leadership traits. The Leadership Population–Population Symmetry Principle is given by the $45°$ line, where $C_{A,t} = \theta_t^{NC}$. Stability is when θ_t^{NC} and $C_{A,t}$ intersect at the 45-degree line, defining the steady state. Thus, the probability of a non-corrupt leadership is concentrated around θ_t^{NC}, the modal societal value, where leadership traits are likely to mirror this societal norm.

An alternative interpretation of the x-axis is that it reflects the proportion of non-corrupt incidents within the population. As this proportion increases, it amplifies contagion effects and habit formation, reinforcing behavioral norms.

If $C_{A,t}$ is above the $45°$ line and it does not align with the underlying population, then over time $C_{A,t}$ will be pushed back to the line, $45°$, and vice versa. You can see there could be multiple steady states in this framework, akin to Sterk (2016). [9]

Figure 1 shows that as societal corruption increases—shifting the population distribution leftward—habit formation, contagion, incentive distortion, and a shrinking pool of honest individuals will combine to reinforce corruption. This implies that the probability of future leadership being more corrupt will increase in this scenario. While in steady state we have:

$$\Delta\theta_t^{NC} = \Delta\theta_t^{C} = 0$$

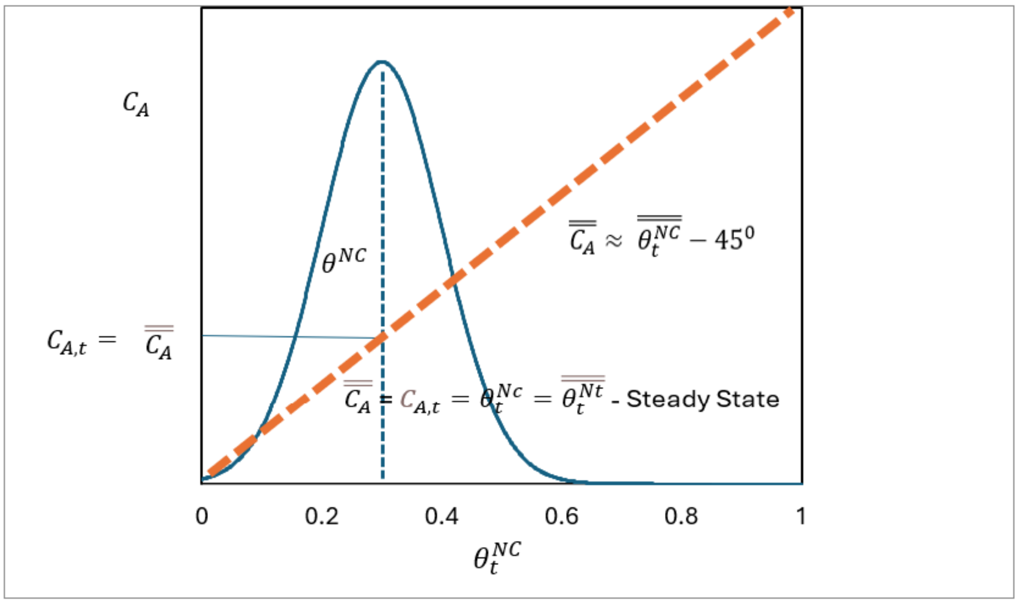


Figure 1: Leadership Population Symmetry Principle

Moving from the model to empirics, this can be stated as the following testable hypothesis.

Hypothesis 1: The more corrupt the social norms of a population, the more likely it is that political leadership will mirror these corrupt norms.

[9] While the core model focuses on the average level of norms in society, the distribution of norms may also matter. For instance, a more dispersed distribution may increase variance in the type of leader selected, while a bi-modal (e.g., polarized) distribution could lead to oscillating leadership types. The U.S. provides a possible example, where both left- and right-leaning leaders are regularly elected. While this is an important extension, modeling the distributional aspects is a paper in itself. Moreover, examining the distribution is not necessary for the core insights of this paper and is left for future research. This framework also illustrates how the model—and the concept of population norms—could be extended to explain leadership outcomes under conditions of polarization driven by rising inequality, social media, or other societal fragmenting forces.

The Role of Leadership Characteristics, C_A

Given the strong influence of societal norms—operating through incentives, contagion, habit formation, selection effects, and normative alignment—meaningful institutional transitions require a shift in these norms from corrupt to non-corrupt. Merely changing laws or formal rules is insufficient. Since societal norms exhibit inertia (North, 1990), such transformation is necessarily slow. This further aligns with North's argument that informal rules and norms often exert greater influence than formal structures.

I argue that, in rare cases, strong and persistent leadership can overcome this inertia and gradually reshape the underlying population characteristics. I further show why this is rare.

In this model, the emergence of leadership is treated as a stochastic process in the long run. In a corrupt society, where the population is centred around a high mean level of corruption, the likelihood is that emerging leaders will also be corrupt. However, since leadership arises from the broader distribution, deviations from this norm can occur.

These deviations allow for the emergence of reformist leaders—such as Lee Kuan Yew in Singapore, Deng Xiaoping in China, or Atatürk in Türkiye—who break from the prevailing societal equilibrium. In these instances, the leadership changed the composition of the population through education, strong judiciary, strong anti-corruption measures including signaling and legislative reforms, and focusing on a merit-based economic system (Ayuso Castillo, 2020; Pamuk, 2007; Morton, 2018; Trevaskes, 2002; Liu and Zhang, 2017; Morton and Ramsay, 1994).

The closest idea to the above, where leadership can on rare occasions transform societies, parallels Acemoglu and Jackson (2015), who suggest that the impact of history can be countered by occasional "prominent" and "exogenous" agents. These individuals, whose actions are highly visible, can shift the expectations of future agents and overturn entrenched social norms. They illustrate this mechanism with the example of Nelson Mandela in South Africa.

When a reformist leader appears, i.e., $C_{A,t} > \theta_t^{NC}$, this represents a shock to the system, and I assume that the leadership will seek to transform the society. As postulated from the symmetry of population and leadership, a meaningful transformation requires the underlying norms on corruption of the population to shift from corrupt to non-corrupt. A central point here is that to make such a meaningful transformation to population characteristics, you need stability and consistency in leadership and reforms. When leadership endures, a *Leadership Hysteresis Effect* can emerge—embedding long-run change that persists beyond the leader's tenure.

I argue that the following two conditions are necessary and sufficient for institutional change:

1. Sustainable institutional change requires changing the characteristics of the population itself.
2. To do this, leadership must exhibit both high reform intensity and sufficient duration—capturing the conditions for the hysteresis effect.

Hypothesis 2: For reforms to have a lasting effect, they need intensity and duration; only then can they lead to a hysteresis effect. Where both intensity and duration are needed to change underlying norms.

Dynamic Interactions Between C_A and Population —Transformation of Norms

This section introduces a difference equation framework showing how sustained leadership can gradually shift societal traits on corruption. When $C_{A,t}$ is high relative to θ_t^{NC}, it gradually gains significance by influencing θ_t^{NC} over time, as shown in Equation (10).

Initial population (P_0) is endowed with a certain proportion of corrupt and non-corrupt population, θ_0^C and θ_0^{NC}, respectively, where the subscript zero refers to the initial period.

$$P_0 = P_0\theta_0^C + P_0\theta_0^{NC} \tag{9}$$

Thereafter the evolution of population is given in Equation (10):

$$P_{t+1} = P_t - \delta_C P_t\theta_t^C - \delta_{NC} P_t\theta_t^{NC} + \mu P_t\theta_t^{NC} + \omega P_t\theta_t^C \tag{10}$$

Since $1 = \theta_t^C + \theta_t^{NC}$, Equation (10) becomes:

$$P_{t+1} = P_t - \delta_C P_t(1 - \theta_t^{NC}) - \delta_{NC} P_t\theta_t^{NC} + \mu P_t\theta_t^{NC} + \omega P_t(1 - \theta_t^{NC}) \tag{11}$$

Where δ_C and δ_{NC} represent the exit rate of corrupt and non-corrupt individuals, respectively. δ_C includes both natural attrition and the removal of corrupt individuals by the judicial system. Hence $\delta_C P_t\theta_t^C$ and $\delta_{NC} P_t\theta_t^{NC}$ give the number of corrupt and non-corrupt people that are taken out of the population.

$\mu P_t\theta_t^{NC}$ and $\omega P_t\theta_t^C$ represent the new population entrants. μ and ω determine the proportion of non-corrupt and corrupt individuals entering the population, respectively.

Equation (10) resembles the logic of capital accumulation equations in macro models, where the composition of the population evolves over time through exit (depreciation) and entry (investment) flows of corrupt and non-corrupt individuals. In this analogy, the "stock" is the population share, and its ethical composition shifts via the "depreciation" of corrupt and non-corrupt individuals and "investment" in honest entrants.

Motivations for ω, μ, δ_C, δ_{NC}

δ_C, ω — When reformist leadership emerges, it can strengthen judicial institutions and increase the probability of punishment, thereby raising the exit rate of corrupt individuals (δ_C) and discouraging corrupt entrants (ω). As suggested in the PA framework, reformist leaders often reshape the judiciary by appointing officials aligned with integrity and enforcement goals. This enhances both formal deterrence (via punishment) and informal norm via signaling.

μ, δ_{NC} — At the same time, a reformist leadership pursuing education reform, strong signaling, and meritocratic appointments will influence the composition of new entrants. These efforts increase the inflow of non-corrupt individuals (μ) and protect honest agents already within the system, thereby reducing the exit rate of non-corrupt individuals (δ_{NC}).

Singapore, Türkiye, South Korea, China, and Botswana exemplify how ω, μ, δ_C, δ_{NC} played a key role in their transitions. In each case, sustained leadership not only enforced legal and institutional reform but also reshaped societal norms through a multi-pronged strategy—strengthening the education system (Ayuso Castillo, 2020; Pamuk, 2007), transforming the legal and judicial framework (Gao and Yao, 2016; Kuru, 2009; Williams, 2014), punishing corruption with credible enforcement mechanisms (Morton, 2018; Trevaskes, 2002), and recruiting based on merit (Liu and Zhang, 2017; Morton and Ramsay, 1994). These actions increased the exit rate of corrupt individuals (δ_C), raised the inflow of non-corrupt individuals (μ), limited (ω) and ultimately increased θ_t^{NC}.

For simplicity, I assume that when $C_{A,t}$ is high, we have $\mu = \sigma$ and $\omega = 0$. Or when $C_{A,t}$ is low we have $\mu = 0$ and $\omega = \sigma$, where σ is some high value and $\sigma < 1$. This means that strong leadership increases the proportion of non-corrupt individuals, and vice versa.

Transitional dynamics

In the long run, the system reaches equilibrium where $C_{A,t} = \theta_t^{NC}$. At this steady state, the exit and entry rates of corrupt and non-corrupt are equal. Under these conditions, θ_t^{NC} becomes constant. As such, in steady state we have:

$$\Delta\theta_t^{NC} = \Delta\theta_t^{C} = 0$$

The model yields several important insights for the transitional dynamics that occur before reaching this equilibrium.

Simplifying Equation (11) we get:

$$P_{t+1} = P_t - (\delta_{NC} - \delta_C)\theta_t^{NC} P_t - \delta_C P_t + (\mu - \omega)\theta_t^{NC} P_t + \omega P_t \tag{12}$$

For simplicity I assume no population growth, such that:

$$g = \frac{P_{t+1} - P_t}{P_t} = 0 \Rightarrow \frac{P_{t+1}}{P_t} = 1$$

$$g = (\mu - \omega)\theta_t^{NC} + \omega - (\delta_{NC} - \delta_C)\theta_t^{NC} - \delta_C \tag{13}$$

$$(\mu - \omega)\theta_t^{NC} + \omega - (\delta_C - \delta_{NC})\theta_t^{NC} - \delta_C = 0 \tag{14}$$

Further, in transition periods we get:

$$\theta_t^{NC} = \frac{\delta_C - \omega}{\mu - \omega + \delta_C - \delta_{NC}} \tag{15}$$

During the transition period, when $C_{A,t} > \theta_t^{NC}$, the system requires the following to increase the share of non-corrupt population:

- A higher inflow of non-corrupt individuals: $(\mu - \omega) > 0$ Eq (16)
- Greater removal of corrupt individuals through enforcement mechanisms: $(\delta_C - \delta_{NC}) > 0$ Eq (17)

In this way, leadership actively drives the transition by increasing μ and δ_C, while minimizing δ_{NC} and ω.

These differences are functions of the distance from equilibrium and the time required to reach it:

$$(\mu - \omega) = f(C_{A,t} - \theta_t^{NC}, T') \tag{18}$$
$$(\delta_{NC} - \delta_C) = f(C_{A,t} - \theta_t^{NC}, T') \tag{19}$$

Where:

- $C_{A,t} - \theta_t^{NC}$ captures the gap between the quality of leadership and the current population norm.

- T' represents the target time horizon to reach steady state where $C_{A,t} = \theta_t^{NC}$.

The larger the gap $C_{A,t} - \theta_t^{NC} > \tau$, the greater the required differential in inflows of non-corrupt individuals ($\mu - \omega$) and the differential in exit rates through the judicial system ($\delta_{NC} - \delta_C$) to achieve convergence within the desired timeframe T'.

Because there are more variables than constraints, I impose additional ones: $0 \leq \theta_t^{NC} \leq 1$, and in the absence of population growth, rising μ implies falling ω. Alternatively, during reform, $(\omega - \delta_C) < 0$ and $(\mu - \delta_{NC}) > 0$, since increasing non-corrupt entry and reducing corrupt entry are part of reform periods.

Graphical illustration

Figure 2 shows this graphically. If $C_{A,t}$ is high relative to θ_t^{NC}, then concerted effort will start to change the underlying population to the right. That is, if $C_{A,t}$ is large enough for long enough, it will shift the underlying population characteristics, shifting $\theta^{NC'}$ to $\theta^{NC''}$ over a period. Once this happens, the future prospect for good leadership increases from $C_A^{'}$ to $C_A^{''}$, as reflected by the underlying population and the symmetry condition.

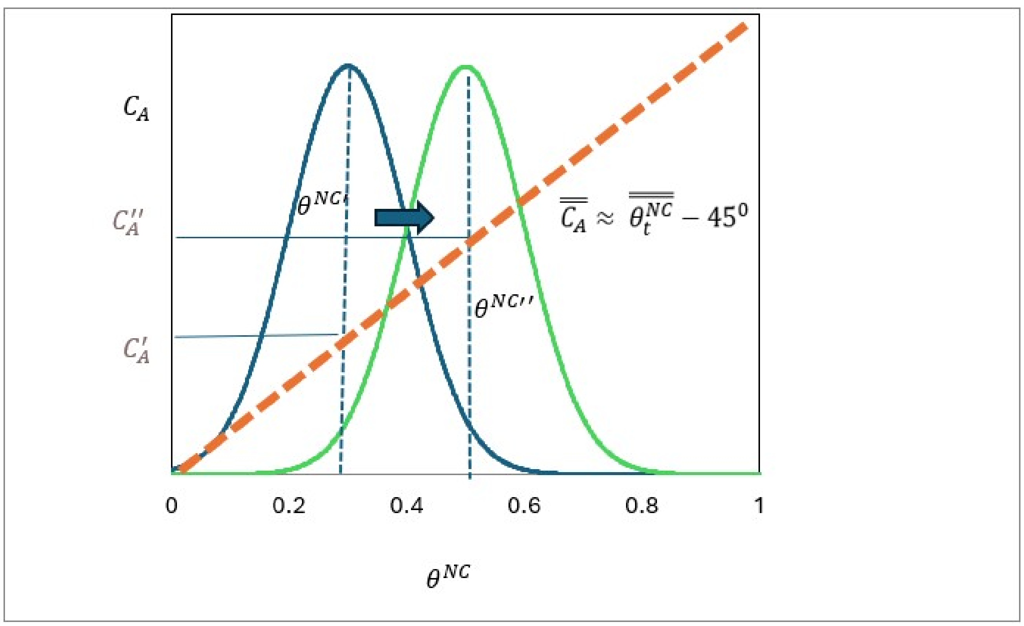


Figure 2: Transition dynamics as leadership reshapes population traits

While Acemoglu and Robinson (2025) emphasize punctuated shifts in institutional equilibria following leadership or political shocks, Figure 2 illustrates that even after a discontinuous change in leadership (i.e., $C_A' \to C_A''$), population norms adjust slowly (from $\theta^{NC'}$ to $\theta^{NC''}$). This reflects a hysteresis dynamic, where reform-minded leaders may push the system toward a new equilibrium, but the speed of convergence depends on structural gaps given by $(\mu - \omega)$, $(\delta_{NC} - \delta_C)$, and the starting point of social norms, $\theta^{NC'}$. Hence, even when formal power changes, informal institutions may take time to follow. [10]

Probability of Change

The model implies that successful institutional change is exceedingly rare due to the compounded improbability of three conditions. First, in a highly corrupt society, the probability of a non-corrupt leadership emerging is low (P_{NC}). Second, even if such leadership arises, designing and implementing intensive and effective reforms is immensely challenging—making the probability of reform intensity (P_i) low. Finally, for these reforms to endure, they must be sustained over an extended period. When $C_{A,t}$ is significantly higher than θ_t^{NC}, leadership will have to continuously push against deeply ingrained societal norms while facing resistance from entrenched elites seeking to preserve the status quo. Thus, the likelihood of reformist leadership maintaining power for sufficient duration (P_{ld}) is also low.

The overall probability of successful institutional change, $P_{success}$, is the product of these probabilities:

$$P_{success} = P_{NC} \cdot P_i \cdot P_{ld} \tag{20}$$

Given that each component is low, their product implies that the probability of successful reform is even smaller.

While the cases of China, South Korea, Singapore, Türkiye, and Botswana are notable examples of successful institutional reform, they remain rare exceptions. The Leadership Hysteresis Effect explains their success, but the broader Population–Leadership Symmetry Principle shows why such cases are uncommon: shifting societal norms is a difficult and long process. The symmetry principle explains the norm—why most societies remain stuck—while the hysteresis effect accounts for rare shifts. As the model formalises the joint occurrence of P_{NC}, P_i and P_{ld} is uncommon, rendering durable reform episodes as exceptions rather than the rule.

Hypothesis 3: Durable institutional transformation is unlikely because it requires the rare convergence of non-corrupt leadership, reform intensity, and political longevity.

[10] The framework echoes Wigner's seminal notion of symmetry in physics, where systems remain stable under specific transformations (Wigner, 1964, 1965). Institutional persistence arises when leadership traits align with prevailing social norms—preserving a form of 'social invariance.' Just as physical systems shift when symmetry is broken, institutional transformation occurs only when this alignment is disrupted by sufficiently intense and sustained reform.

Causality

Leadership typically reflects prevailing societal norms; the symmetry principle thus implies a directional relationship from societal norms to leadership traits. Yet, in rare cases, leadership can reshape those very norms. The Leadership Hysteresis Effect outlines the mechanisms by which sustained and intensive leadership can drive such transformation in rare cases.

Threshold Effect

I also introduce a threshold effect: once θ_t^{NC} exceeds a critical value $\hat{\theta}_t^{NC}$, it becomes difficult for $C_{A,t}$ to reverse institutional quality. At this point, not only does a high θ_t^{NC} mean that most policymakers and leaders are likely to be non-corrupt, but it also ensures that there is a sufficiently large group of citizens to actively resist against backsliding and corruption in leadership. This includes applying public pressure—such as protests—to uphold institutional integrity. [11]

5 Calibration to Reform-Era Transitions

This section calibrates the model to the average reform trajectory across four countries—Singapore, South Korea, Türkiye, and Botswana—during periods of sustained, high-intensity reforms aimed at reducing corruption. They serve as benchmark cases for evaluating the model's predictions under rare conditions of successful, sustained reform.

Direct measures of societal corruption norms are unavailable over long historical periods. As such I will use a proxy based on public sector corruption that is available.

At the outset, the average share of non-corrupt individuals in the pubic was approximately 0.57, based on V-Dem's 'Public Sector Corrupt Exchanges' indicator. This gradually increased to a final value of 0.77, reflecting a substantial shift in corruption norms over 24 years in these four countries. Over the same period, the average population rose from 40.4 million to 62.9 million. Table 1 presents the full set of calibrated values for $\bar{\omega}$, $\bar{\mu}$, $\bar{\delta}_C$, and $\bar{\delta}_{NC}$.

[11] An example of this is South Korea, where the President attempted to impose martial law, but it resulted in mass protests by the people.

Table 1: Calibrated Parameter Values

Parameter	Initial Value	Reform Years (0–24)	Post-Reform Value	Actual	Calibrated Final Values
θ_t^{NC} (non-corrupt share)	0.565	Evolves via transition equation	0.83		∼0.80
θ_t^{C} (corrupt share)	0.435	$1 - \theta_t^{NC}$	0.17		∼0.20
P_0 (population, millions)	40.4	Linear growth to 62.9M	62.9M		62.9
μ (non-corrupt entry)—early	0.045	Increases gradually from 0.045	—		—
μ (non-corrupt entry)—late	0.095	Peaks at 0.095 in Year 24	—		—
ω (corrupt entry)	—	0.02	0.02		0.02
δ_{NC} (exit of non-corrupt)	0.02	0.02	0.02		0.02
δ_C (exit of corrupt)	—	0.11	0.02		0.02

By adjusting key parameters—such as increasing the inflow of non-corrupt individuals and the exit rate of corrupt individuals (δ_C), and reducing the attrition of non-corrupt individuals (δ_{NC}) and entry of ω—the model closely matches the observed rise in the non-corrupt share. These calibrated trajectories mirror the documented decline in corruption across the four economies during their reform periods.

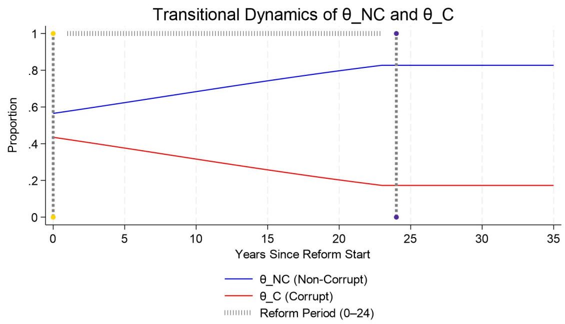
The parameter μ_t (entry of non-corrupt individuals) evolves gradually during the reform period, increasing linearly from 0.045 in year 0 to a peak of 0.095 by year 24. I assume:

$$\mu_t = 0.045 + 24(0.095 - 0.045) \cdot t \tag{1}$$

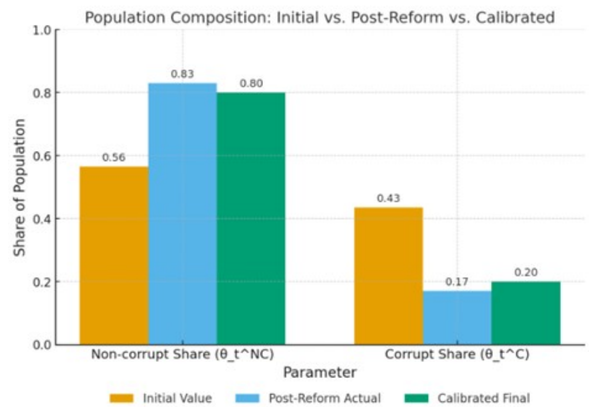
where t is the number of years since reform began. This formulation captures both habit formation and the lagged effects of educational reforms that gradually influence the composition of new entrants.

Figure 3.2, Panels A and B, show that the model produces the empirical path of corruption with close fit. This suggests that the model not only captures the key dynamics of institutional transformation but also offers a tractable framework for interpreting historical episodes and guiding future reform strategies.

The calibration findings—and the actual evidence from these countries—reinforce a core insight of the model: transforming population norms is inherently gradual and requires sustained reformist leadership, consistent with the dynamics of the leadership hysteresis effect.



(a) Panel A: Model Calibration — Share of Non-Corrupt Population



(b) Panel B: Model Calibration Values

Figure 3: Calibration

6 Tests of the model

I now empirically test the model's two core hypotheses.[12]

I begin by testing Hypothesis 1. A key limitation in testing this claim is the absence of a valid external instrument, which constrains causal interpretation. Accordingly, I interpret the results as evidence of association rather than causation. Nevertheless, I test Hypothesis 1 across a wide range of methodologies and samples, and find consistent and suggestive evidence in its favor.

The empirical analysis primarily focuses on Hypothesis 1, but I also provide initial evidence for Hypothesis 2. To test this, I examine reform episodes across countries—contrasting settings where leadership was both strong and sustained (e.g., Singapore under Lee Kuan Yew or Türkiye under Atatürk) with those where leadership was strong but relatively short-lived (e.g., Pakistan under Pervez Musharraf). These comparisons allow me to assess the presence of a leadership hysteresis effect, whereby reforms become self-reinforcing over time.

Data

This section briefly outlines the data sources used in the analysis. When bringing the prediction of the model to the data, a key issue is how to measure corruption norms in the society versus corruption norms in the leadership. A central empirical challenge here is that standard corruption indices often conflate leadership behavior with societal

[12]To keep the paper concise, I leave Hypothesis 3 for future research.

norms. To address this, I rely on individual-level data from the WVS to capture ethical attitudes toward corruption among individuals and Worldwide Governance Indicators (WGI) indicators to capture the leadership norms.

From WVS I construct country-level indicators of ethical tolerance, averaging individual responses to key corruption-related questions. The main variables measure whether respondents believe it is justifiable to (i) claim government benefits to which they are not entitled (Q177), (ii) not pay fare on public transport, and (iii) cheat on taxes (Q180)[13]. I also take a mean of these values to create an average indicator. These questions reflect public attitudes toward actions that undermine integrity and accountability. From the above WVS questions I am able to identify an individuals norms and preferences on corruption, and by aggregating them to a country index, I am able to create an estimate of society's norm on corruption.

To test the Population–Leadership Symmetry Principle, the dataset spans 1995 to 2020, ensuring that WGI and WVS coverage aligns. The analysis draws on a broad country sample, with the main panel regressions restricted to 44 countries with at least two time observations to ensure within-country variation, while pooled OLS specifications include up to 93 countries (see Appendix for full list). As WVS data is collected in waves, I link each wave to the corresponding WGI year, resulting in an average four-year gap between observations per country.

The WVS comprises five waves (1996–2020), with each country typically contributing between 7,000 and 10,000 observations. While the broadest regressions incorporate all available waves from 1996 onward, the main analysis begins in 2007. This cutoff reflects the limited coverage and inconsistent sampling of earlier waves (e.g., 1982–2002), which include fewer than 30 countries per year on average. From 2007 onward, the WVS sample becomes more stable and comprehensive, with approximately 60 or more countries per wave, ensuring stronger cross-country comparability.

To examine government leadership, I rely on the Control of Corruption indicator from the WGI project. This metric captures the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as state capture by elites. It closely reflects the ethical dimension of leadership emphasized in Hypothesis 1. As a robustness check, I also use the Rule of Law indicator from WGI. While it captures aspects related to leadership, such as public confidence in the judiciary and enforcement credibility—it is broader in scope, encompassing crime rates, judicial efficiency, and contract enforcement.[14]

[13]I exclude the variable on taking bribes due to conceptual and measurement concerns. Unlike the other questions, which ask about personal dishonest behavior, this item refers to bribe-taking which is typically done by public officials. As a result, respondents may interpret it inconsistently—some may view it as justifying their own acceptance of bribes, while others may see it as referring to officials extracting bribes from them

[14]Similarly, other WGI indicators, such as Government Effectiveness, Regulatory Quality, Political Stability, and Voice and Accountability, include dimensions related to leadership qualities. However,

While these variables help distinguish between leadership and population corruption, the sample size poses challenges. WGI data begins in 1996, and WVS waves occur approximately every four years, limiting the number of observations. Nevertheless, the dataset retains sufficient variation for meaningful analysis. To account for this small sample constraint, I apply finite-sample corrections in the Generalized Method of Moments (GMM) estimates.

While WGI indicators are widely used to capture institutional quality, I also assess the model's predictions using an alternative source: the V-Dem dataset. This provides a complementary test, drawing on measures of particularistic spending (v2dlencmps), which captures bias in public goods provision, and executive bribery (v2exbribe), which reflects high-level illicit exchanges within the leadership.

Table 2: Summary Statistics

Variable	**Source**	**N**	**Mean**	**SD**	**Min**	**Max**
Education Expenditure (% GDP)	WDI	187	4.63	1.45	1.75	9.90
Trade Openness (% GDP)	WDI	238	74.81	58.02	15.64	425.98
Log of GDP per capita	WDI	240	9.53	1.00	6.90	11.60
Justifiable: Gov't Benefits (Q177)	WVS	248	2.70	0.88	1.34	6.18
Justifiable: Fare Evasion (Q178)	WVS	247	2.64	0.86	1.13	5.76
Justifiable: Tax Cheating (Q180)	WVS	245	2.28	0.70	1.00	5.67
Control of Corruption	WGI	248	0.12	1.09	-1.60	2.38
Rule of Law	WGI	248	0.14	1.03	-2.33	2.00
Government Effectiveness	WGI	247	0.27	0.96	-2.09	2.32
Political Stability	WGI	247	-0.14	0.96	-3.18	1.62
Regulatory Quality	WGI	247	0.25	1.00	-2.25	2.23
Voice & Accountability	WGI	248	0.10	0.95	-1.91	1.74
Public Sector Corrupt Exchanges	V-Dem	242	2.06	1.01	0.14	4.00
Public Sector Theft	V-Dem	242	2.46	0.97	0.07	4.00
Freedom of Expression	V-Dem	242	0.72	0.26	0.05	0.98
Electoral Democracy Index	V-Dem	242	0.58	0.26	0.02	0.92

Data sources: WDI — World Development Indicators; WGI — Worldwide Governance Indicators; WVS — World Values Survey.

Hypothesis 1: Estimation

I estimate the following three equations:

these indicators extend beyond leadership integrity: for example, Government Effectiveness reflects bureaucratic quality, Regulatory Quality concerns market regulation, Political Stability captures risks of violence, and Voice and Accountability measures civil liberties and media freedom. While valuable for some measure of robustness, these are less precise proxies for the ethical dimension of leadership that is central to this paper.

$$Leadership\_characteristics_{it} = \beta_0 + \beta_1\, SocialNorm_{it} + \boldsymbol{\beta}\,\mathbf{X}_{it} + \delta_t + \varepsilon_{it} \tag{E1}$$

$$Leadership\_characteristics_{it} = \beta_0 + \beta_1\, InitialSocNorm_i + \boldsymbol{\beta}\,\mathbf{X}_{it} + \delta_t + \varepsilon_{it} \tag{E2}$$

$$Leadership\_characteristics_{it} = \beta_0 + \beta_1\, InitialSocNorm_i + \beta_2\, \Delta SocNorm_{it} + \boldsymbol{\beta}\,\mathbf{X}_{it} + \delta_t + \varepsilon_{it} \tag{E3}$$

Here, Norm captures each country's integrity population norms, while ΔSocNorm measures changes. **X** is a vector of country and regional-level controls, described below in Table 3. δ_t are time dummies and ε_{it} is the error term.

Two key challenges to Equation (E1) are reverse causality—where leadership may shape societal values—and omitted variable bias, where omitted factors could jointly influence both leadership characteristics and societal norms. As such, the robustness tests would control for such confounding factors.

X is designed to capture variables that might be correlated with social norms but may independently impact the level of corruption in the leadership. For instance, increased integration with the global economy can diffuse new norms and best practices, improving both governance and societal attitudes. To account for this, I control for trade openness (trade as a share of GDP) as a proxy for international exposure. Similarly, higher levels of education may simultaneously raise ethical standards among the population and enhance leadership capacity, so I include controls for education levels. I also include contemporary values of the natural log of a country's real per-capita GDP. This captures differences in economic development, which could affect leadership and population norms through channels other than the one I am interested in identifying.

Robustness checks incorporate additional covariates, including democracy, freedom of expression, and legal origins and regional trends. These variables are discussed in the robustness sections.

Table 3: Different combinations of determinants of leadership characteristics (**X**) used in empirical specification

	Set 1	Set 2	Set 3	Set 4	Set 5
Education expenditure (% GDP)	X	X	X	X	X
Trade openness (Trade as % of GDP)	X	X	X	X	X
GDP per capita (log)		X	X	X	X
Democracy Index			X		X
Freedom of expression Index			X		X
Legal origins				X	X
Regional effects				X	X

To mitigate reverse causality concerns, I adopt a temporal separation strategy using Equations (E2) and (E3). Here, *Initial SocNorm* captures each country's lagged level of corrupt population norms. Specifically, I use lagged values of societal norms—measured 5 to 20 years prior—to examine associations with subsequent institutional quality. This approach assumes that earlier values of population attitudes are unlikely to be shaped by current governance performance, allowing for a more credible assessment of directional influence from norms to institutions Breunig and Majeed (2020); Cingano (2014). Additionally, I control for changes in norms (ΔSocNorm) in Equation (E3). The main challenge with Equation (E3) is it drops countries that does not have consecutive observations. However, results remain robust to both versions.

First difference and system GMM techniques overcome reverse causality and omitted variable biases as well, including controlling for country fixed effects. First difference GMM remedies these problems by taking the first difference of the equation to remove country fixed effects and using appropriately lagged values of dependent and explanatory variables as internal instruments. However, the first difference transformation suffers from the problem of weak instruments if the right-hand side variables are highly persistent, which is likely to be the case for societal norms and education, as recognised by Halter et al. (2014).

System GMM overcomes this problem by building a system of level and first difference equations and using appropriately lagged instruments, following Arellano and Bond (1991); Roodman (2009); Breunig and Majeed (2020). Further, the first difference methodology has the problem of magnifying issues in unbalanced panels, so instead I use orthogonal deviations, constructed as in Roodman (2009).

I apply the Windmeijer (2005) finite-sample correction to improve the reliability of standard errors in small samples. Initial values of population norms are included as predetermined covariates, under the assumption that future leadership quality cannot influence past societal attitudes. This allows identification of the effect of social norms on the evolution of leadership quality. For robustness, all regressions are done in both OLS and GMM estimations. [15]

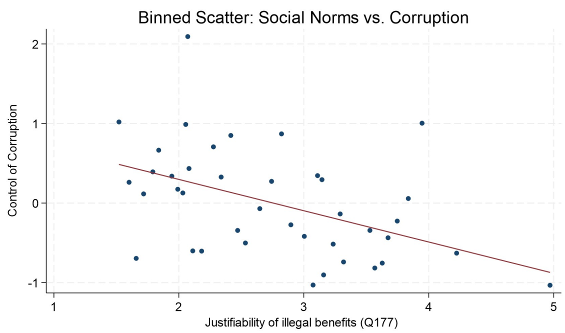
The main empirical limitation of this paper is the absence of a valid external instrument, which constrains the ability to draw causal inferences. As emphasized in Acemoglu and Robinson (2025) establishing causality in this domain is inherently challenging. Accordingly, the results should be interpreted as evidence of associations rather than definitive causation.

Results

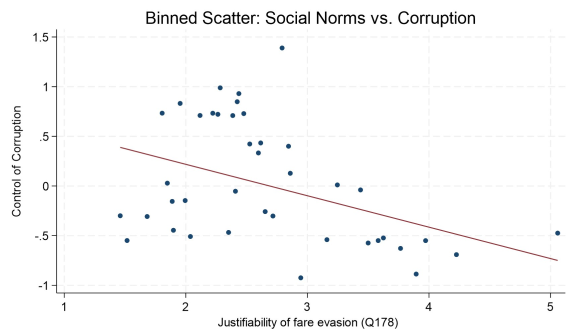
OLS

[15] For the main regressions, I drop all countries with fewer than two observations to ensure sufficient within-country variation over time.

Figure 3 presents the bivariate relationship between societal norms and leadership corruption. Panels A and B shows a simple cross-country scatter plot between the belief that it is justifiable to claim benefits dishonestly and fare evasion with Control of Corruption, revealing a clear negative association. This appears to be general and not driven by outliers; removing outliers (95th percentile of observations) from the regressions does not change the results.



(a) Panel A: Justifiable to take illegal benefits



(b) Panel B: Justifiable to evade fares

Figure 4: Association Between Societal Beliefs and Leadership Integrity—Binned Scatter

Tables 4 and A2 report OLS estimates of the relationship, using Control of Corruption and Rule of Law as dependent variables, while using four societal values for claiming as variants of population characteristics. In odd-numbered columns, I use no controls, while in even-numbered columns, I use control set 2. Control Set 1 yields qualitatively similar results [16], but as control set 2 additionally controls for level of development—which is a key variable—this is my preferred control set.

I find a negative and statistically significant relationship: societies with greater behavioral tendencies to cheat are associated with weaker leadership integrity.

Examining the coefficient estimates for the additional control variables, we see that economic characteristics—such as education expenditure, trade openness, and income—are also associated with leadership integrity, though not always statistically significant. In particular, higher education spending and GDP per capita are generally correlated with improvements in the control of corruption. While trade openness is statistically significant, it is not economically meaningful. The positive association between income and control of corruption is also consistent with the idea that higher-income societies may have stronger governance institutions or a higher demand for integrity in leadership Giuliano and Nunn (2021).

Results remain qualitatively similar for four variables on societal norms. However,

[16] Results available on request.

as mentioned earlier, societal values on taking bribes are likely to exhibit measurement error, and as such I drop it from further analysis. Results are also qualitatively similar if I use other WGI variables; however, as discussed earlier, these variables are not of primary interest for this study. As such, the rest of the results do not use them [17].

OLS estimates from Equation (E1) may suffer from reverse causality and omitted variable bias, though the control set helps mitigate these concerns. The following subsections address these issues using lag structures, GMM, and additional controls, with results remaining robust throughout.

Managing Reverse Causality

$$LastValue\_LeadershipCharacteristics_i = \beta_0 + \beta_1\, InitialSocNorm_i + \boldsymbol{\beta}\,\bar{\mathbf{X}}_i + \varepsilon_i \qquad \text{(E4)}$$

To assess whether the main findings are driven by reverse causality, I estimate a set of OLS regressions where the final observed values of leadership characteristics (from 2020) are regressed on the initial values of societal norms. Country-level averages of Control Set 2 variables are included as covariates. I use the 2007 as the baseline for initial societal norms, given its substantially broader country coverage than earlier waves.

This strategy exploits a 13-year temporal gap between the dependent and independent variables, mitigating concerns about potential reverse causality Breunig and Majeed (2020); Cingano (2014): current leadership characteristics are unlikely to retrospectively influence past societal beliefs about cheating and integrity. As shown in Figure 5, the results remain robust to this alternative specification. Covariates are averaged across available years for each country. Full results are presented in Table 5.

The results show a clear negative and statistically significant relationship: greater societal tolerance for cheating is associated with lower subsequent leadership quality.

The estimated effect is also economically meaningful. For example, in Table 5, Column (2), a one-standard-deviation increase in societal tolerance for claiming illegal benefits (0.87) is associated with a $0.43 \times 0.87 \approx 0.38$ point decline in the control of corruption index. Given that the corruption index has a standard deviation of approximately 1.0, this corresponds to a drop of about 38% of a standard deviation—a sizable effect.[18]

Robustness: Results remain consistent when varying the starting year based on each country's first available observation—beginning as early as 1996—thereby generating temporal gaps of 8 to 25 years between initial societal norms and final observed

[17]Results available on request.

[18]If I relax the requirement that countries should at least have two observations of societal values, then sample of countries goes to around 56 and results still remain robust.

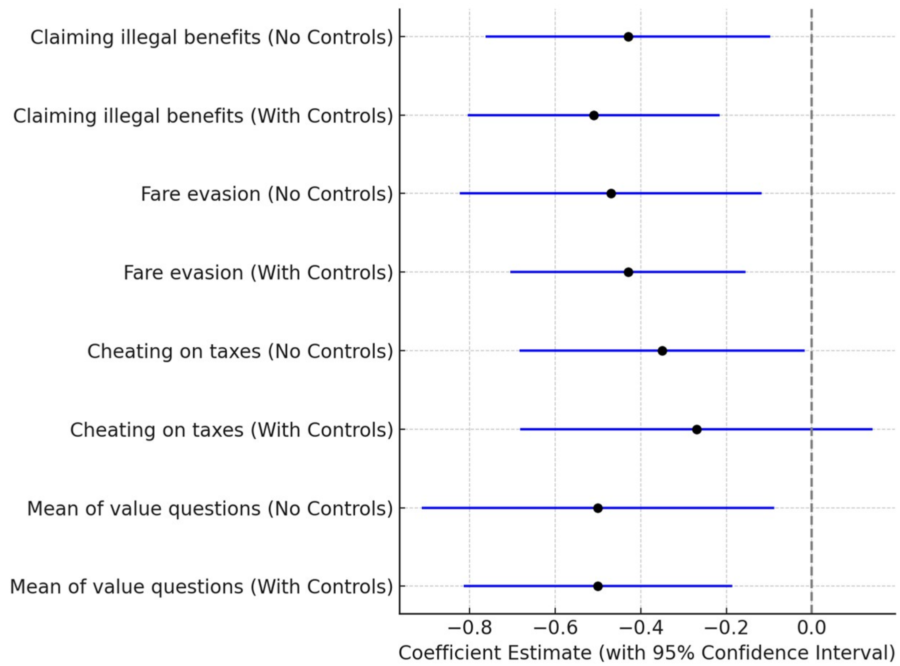


Figure 5: Effect of Initial Societal Norms on Last Leadership Characteristics

Note: Control variables (Control Set 2) are averaged for each country. Mean value includes means of social value questions on tax, fare evasion, and illegal benefits.

leadership outcomes. This allows regressions to have around 95 to 105 countries in the regressions. The findings are also robust to using Rule of Law, instead of Control of Corruption, as the dependent variable.

Baseline regressions with OLS without change over time.

The previous section showed that the results remain robust even when addressing reverse causality by introducing long lags. This section turns to Equation (E2). Table 6 reports the results. Again, across all columns, weaker societal norms—measured by greater tolerance for cheating—are associated with a negative and statistically significant relationship with leadership quality.

Again, not only are the estimated coefficient for societal tolerance of corrupt behavior statistically significant, they are also economically meaningful. For instance,based on the coefficient on claiming illegal government benefits in Table 6, Column (2), a one-standard-deviation increase in societal tolerance for claiming illegal government benefits (0.869) is associated with a decline of $0.40 \times 0.869 \approx 0.35$ in the control of corruption index. Given that the standard deviation of the corruption index is approximately 1.00, this corresponds to a decline of roughly 35% of a standard deviation in leadership quality—a substantial effect. Similar economic magnitudes are observed for the remaining norms variables across the following tables; calculations are omitted for conciseness.

OLS regression—Equation (E3).

Table 7 reports results that additionally control for change in social norms from the previous period (Equation (E3)). Similar to previous results, across all columns, weaker societal norms—measured by greater tolerance for cheating—are associated with a negative and statistically significant relationship with leadership quality. Adding change of consecutive periods does not alter the results qualitatively, though the number of observations drops.

Internal instruments for reverse causality and country fixed effects — GMM

$$Leadership\_characteristics_{it} = \beta_0 + \beta_1\, InitialSocNorm_i + \beta_2\, \Delta SocNorm_{it} + \boldsymbol{\beta}\, \mathbf{X}_{it} + \delta_t + \mu_i + \varepsilon_{it} \tag{E5}$$

An additional robustness test is to estimate System GMM models, controlling for reverse causality and omitted variables bias, using Equation (E5). δ_t are time dummies,

μ_i is country fixed effects, and ε_{it} is the error term. This approach instruments for regressors using their own lagged values. Tables 8 and 9 present GMM results with and without change in social values.

Similar to OLS results, I find a negative and statistically significant relationship: initial societal values with greater behavioral tendencies to cheat are associated with weaker leadership integrity. Results remain robust.

Robustness. All GMM specifications use one lag of internal instruments. The results remain robust when using two lags instead (available on request). Since OLS and GMM yield similar estimates—and including changes in societal norms across periods does not alter the core findings—I use GMM estimates using Equation (E5) for the remaining robustness.

Democracy, Media Access, Legal Origins of Institutions, and Common Institutional Influences

A potential concern is that both population norms and leadership quality may be shaped by broader institutional trends. For instance, the diffusion of democratic ideals or the rise of global media may simultaneously enhance citizen awareness and political accountability. Similarly, regional shocks—such as the Arab Spring or ASEAN's economic integration—and historical factors like legal origin may jointly influence population and leadership norms, introducing omitted variable bias if not accounted for.

To address these concerns, I incorporate controls for institutional diffusion and regional trends into the preferred GMM specification (starting with Set 3). Table A3 presents results controlling for democratic institutions using the Electoral Democracy Index (v2x_polyarchy) and the Freedom of Expression Index (v2x_freexp_altinf) from the V-Dem dataset. These variables proxy for citizens' access to political information and responsiveness of government. Results show that the coefficients on initial societal norms on cheating remain negative and statistically significant, even after controlling for these institutional features.

However, dynamic panel models face the risk of instrument proliferation. When both democracy and freedom of expression are included, the Hansen test statistic rises to 0.98, suggesting weak instrument validity Roodman (2009). As such, while the core results remain robust, I interpret specifications with both democratic controls as supplementary.

Another potential threat to the specification is legal origin. La Porta et al. (2008) show that legal traditions—such as common law or civil law—shape institutional quality and economic development. Common law systems (e.g., English origin) tend to offer stronger investor protections and more adaptable legal processes than civil law systems (e.g., French or German origin). The results remain robust to controlling for legal origin and are available on request.

Another concern is that regional forces—such as the Arab Spring (widespread public mobilization) or ASEAN's rise (regional economic progress)—may jointly influence leadership and societal norms.

To address this, I construct region-year averages of institutional outcomes—excluding each country's own values—based on a regional classification of countries [19]. These regional trends are then included as controls. Formally, for a country i in region r and year t, I calculate:

$$\bar{Y}_t^{(-i)} = \frac{1}{N_{rt} - 1} \sum_{\substack{j \in r \\ j \neq i}} Y_{jt}$$

where Y_{jt} is the value of the outcome variable (e.g., control of corruption or rule of law), and N_{rt} is the number of countries in region r at time t. The results remain qualitatively similar and are available on request.

Alternative Measures of Leadership Quality

As an additional robustness test, I replace the dependent variables from the WGI with leadership characteristic variables from the V-Dem dataset. Specifically, I use Particularistic or Public Goods Provision (v2dlencmps), which captures the extent to which government spending disproportionately benefits specific groups rather than the general public, and Executive Bribery and Corrupt Exchanges (v2exbribe), which measures the degree to which high-level public officials engage in bribery and illicit transactions. Results are presented in Table A4, which reports system GMM estimates using covariate Set 2. Across all specifications, the core results remain robust.

Results Summary

The estimates suggest that societal values—tolerance for cheating—are negatively associated with leadership quality. These findings hold across both OLS and system GMM specifications and remain robust after accounting for potential reverse causality, time-invariant country characteristics, and a broad set of covariates. The results are not only statistically significant but also economically meaningful.

These findings are consistent with Hypothesis 1: the more corrupt the social norms of a population, the more likely that political leadership will mirror these norms.

[19]The regions include: Western Europe, Eastern Europe, South Asia, Arab countries, Rest of Asia, North Africa, Southern Africa, South America, North America, Pacific, and Other.

Table 4: Determinants of Corrupt Leadership: OLS Estimates

VARIABLE	**(1) Con. Crpt**	**(2) Con. Crpt**	**(3) Con. Crpt**	**(4) Con. Crpt**	**(5) Con. Crpt**	**(6) Con. Crpt**	**(7) Con. Crpt**	**(8) Con. Crpt**
Societal values:								
Claiming illegal benefits	**-0.36*****	**-0.24*****						
	(0.08)	(0.08)						
Cheating on taxes			**-0.40*****	**-0.28*****				
			(0.11)	(0.09)				
Fare evasion					**-0.34*****	**-0.26*****		
					(0.09)	(0.07)		
Taking bribes							**-0.42*****	**-0.16***
							(0.11)	(0.10)
Education Expenditure		0.09*		0.10*		0.11**		0.09*
		(0.05)		(0.05)		(0.05)		(0.05)
Trade Openness		0.00		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)
Log GDP per capita		0.80***		0.81***		0.83***		0.83***
		(0.08)		(0.08)		(0.08)		(0.08)
Tax revenue		0.01		0.02*		0.02		0.02*
		(0.01)		(0.01)		(0.01)		(0.01)
Constant	1.40***	-7.36***	1.37***	-7.61***	1.37***	-7.72***	1.21***	-8.10***
	(0.25)	(0.84)	(0.28)	(0.79)	(0.27)	(0.76)	(0.26)	(0.84)
Observations	178	115	176	114	175	113	179	115
R-squared	0.14	0.68	0.12	0.68	0.12	0.69	0.10	0.66
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The sign of the constant reverses when controlling for GDP per capita, reflecting the lower corruption scores of poorer countries.

Table 5: Determinants of Corrupt Leadership: Reverse Causality (OLS)

VARIABLE	**(1) Lt. CCPT**	**(2) Lt. CCPT**	**(3) Lt. CCPT**	**(4) Lt. CCPT**	**(5) Lt. CCPT**	**(6) Lt. CCPT**	**(7) Lt. CCPT**	**(8) Lt. CCPT**
Initial Societal Value:								
Claiming illegal benefits	**-0.51*****	**-0.43*****						
	(0.17)	(0.15)						
Fare evasion			**-0.43****	**-0.47*****				
			(0.18)	(0.14)				
Cheating on taxes					**-0.27**	**-0.35****		
					(0.21)	(0.17)		
Mean of integrity values							**-0.50****	**-0.50*****
							(0.21)	(0.16)
Trade Openness (% GDP)		0.00		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)
Log GDP per capita		0.51***		0.56***		0.57***		0.55***
		(0.13)		(0.12)		(0.13)		(0.12)
Constant	1.71***	-3.93***	1.52***	-4.62***	0.96*	-5.14***	1.62***	-4.39***
	(0.47)	(1.31)	(0.50)	(1.18)	(0.49)	(1.25)	(0.53)	(claims)
Observations	44	41	43	40	44	41	44	41
R-squared	0.18	0.51	0.12	0.54	0.04	0.46	0.12	0.52
Time Effects	No	No	No	No	No	No	No	No
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Determinants of Corrupt Leadership: OLS (Initial Societal Values)

VARIABLE	**(1)** **Con. Crpt**	**(2)** **Con. Crpt**	**(3)** **Con. Crpt**	**(4)** **Con. Crpt**	**(5)** **Con. Crpt**	**(6)** **Con. Crpt**	**(7)** **Con. Crpt**	**(8)** **Con. Crpt**
Initial Societal Value:								
Claiming illegal benefits	**-0.50*****	**-0.40*****						
	(0.10)	(0.08)						
Fare evasion			**-0.37*****	**-0.42*****				
			(0.11)	(0.09)				
Cheating on taxes					**-0.26****	**-0.34*****		
					(0.13)	(0.10)		
Mean of integrity values							**-0.46*****	**-0.47*****
							(0.12)	(0.10)
Education Expenditure		0.14**		0.22***		0.19***		0.19***
		(0.06)		(0.06)		(0.06)		(0.06)
Trade Openness		0.00*		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)
Log GDP per capita		0.66***		0.71***		0.74***		0.70***
		(0.08)		(0.08)		(0.08)		(0.08)
Constant	1.71***	-5.61***	1.40***	-6.37***	0.98***	-6.86***	1.57***	-6.10***
	(0.30)	(0.85)	(0.33)	(0.78)	(0.32)	(0.80)	(0.34)	(0.80)
Observations	132	109	129	106	132	109	132	109
R-squared	0.15	0.61	0.08	0.61	0.03	0.57	0.10	0.61
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Determinants of Corrupt Leadership: OLS with Change in Norms

VARIABLE	(1) Con. Crpt	(2) Con. Crpt	(3) Con. Crpt	(4) Con. Crpt	(5) Con. Crpt	(6) Con. Crpt	(7) Con. Crpt	(8) Con. Crpt
Initial: Claiming illegal benefits	**-0.50*****	**-0.36*****						
	(0.10)	(0.13)						
Change: Claiming illegal benefits		**-0.18**						
		(0.11)						
Initial: Fare Evasion			**-0.37*****	**-0.53*****				
			(0.11)	(0.13)				
Change: Fare Evasion				**-0.30***				
				(0.16)				
Initial: Cheating on Taxes					**-0.26****	**-0.46****		
					(0.13)	(0.17)		
Change: Cheating on Taxes						**-0.29**		
						(0.22)		
Initial: Mean of Integrity Values							**-0.46*****	**-0.54*****
							(0.12)	(0.15)
Change: Mean of Integrity Values								**-0.30***
								(0.16)
Education Expenditure		0.13		0.23***		0.18**		0.18**
		(0.08)		(0.08)		(0.08)		(0.08)
Trade Openness		0.00		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)
Log GDP per capita		0.72***		0.75***		0.71***		0.72***
		(0.14)		(0.13)		(0.14)		(0.13)
Constant	1.71***	-6.40***	1.40***	-6.72***	0.98***	-6.54***	1.57***	-6.30***
	(0.30)	(1.44)	(0.33)	(1.31)	(0.32)	(1.40)	(0.34)	(1.33)
Observations	132	55	129	52	132	53	132	55
R-squared	0.15	0.56	0.08	0.62	0.03	0.53	0.10	0.60
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Determinants of Corrupt Leadership: System GMM Estimates

VARIABLE	**(1)** **Con. Crpt**	**(2)** **Con. Crpt**	**(3)** **Con. Crpt**	**(4)** **Con. Crpt**	**(5)** **Con. Crpt**	**(6)** **Con. Crpt**	**(7)** **Con. Crpt**	**(8)** **Con. Crpt**
Initial: Claiming illegal benefits	**-0.50***** (0.18)	**-0.33**** (0.13)						
Initial: Fare Evasion			**-0.37**** (0.17)	**-0.42***** (0.14)				
Initial: Cheating on Taxes					**-0.26*** (0.14)	**-0.36***** (0.13)		
Initial: Mean of Integrity Values							**-0.46**** (0.18)	**-0.45***** (0.14)
Education Expenditure		0.12 (0.07)		0.17** (0.07)		0.16** (0.07)		0.15* (0.08)
Trade Openness		0.00 (0.00)		0.00 (0.00)		0.00 (0.00)		0.00 (0.00)
Log GDP per capita		0.75*** (0.12)		0.77*** (0.13)		0.73*** (0.12)		0.76*** (0.12)
Constant	1.60*** (0.35)	-6.20*** (1.25)	1.50*** (0.38)	-6.30*** (1.28)	1.00*** (0.36)	-6.10*** (1.24)	1.55*** (0.37)	-6.00*** (1.23)
Observations	132	109	129	106	132	109	132	109
Instruments	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications use one lag of internal instruments.

Table 9: Determinants of Corrupt Leadership: System GMM with Change in Norms

VARIABLE	(1) Con. Crpt	(2) Con. Crpt	(3) Con. Crpt	(4) Con. Crpt	(5) Con. Crpt	(6) Con. Crpt	(7) Con. Crpt	(8) Con. Crpt
Initial: Claiming illegal benefits	**-0.55*****	**-0.28***						
	(0.17)	(0.16)						
Change: Claiming illegal benefits		**0.07**						
		(0.11)						
Initial: Fare Evasion			**-0.45*****	**-0.50****				
			(0.17)	(0.20)				
Change: Fare Evasion				**0.03**				
				(0.25)				
Initial: Cheating on Taxes					**-0.35****	**-0.43****		
					(0.14)	(0.19)		
Change: Cheating on Taxes						**0.08**		
						(0.32)		
Initial: Mean of Integrity Values							**-0.56*****	**-0.50****
							(0.17)	(0.19)
Change: Mean of Integrity Values								**0.17**
								(not reported
Constant	1.82***	-8.77***	1.59***	-6.07**	1.18***	-5.79**	1.79***	-6.57**
	(0.49)	(3.16)	(0.52)	(2.50)	(0.40)	(2.78)	(0.51)	(2.71)
Control Set 2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	220	70	216	66	220	68	220	70
Countries	55	38	54	37	55	37	55	38

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications use one lag of internal instruments. Control coefficients omitted for brevit

Hypothesis 2

Hypothesis (H2): For reforms to have a lasting effect, they need intensity and duration; only then can they lead to a hysteresis effect. Where both intensity and duration are needed to change underlying norms.

The previous section showed that societal norms around cheating are strongly associated with leadership traits across countries—supporting Hypothesis 1, i.e., the Population–Leadership Symmetry Principle based on a broad set of countries. Yet, as formalised in the Leadership Hysteresis Effect and the section on the probability of change, rare but durable institutional shifts can occur when reformist leadership is both intense and sustained. This section turns to those exceptional outlier cases, providing an initial empirical support for Hypothesis 2. The empirical analysis section main focus is on Hypothesis 1, with preliminary support also shown for Hypothesis 2.

I compare Türkiye, Singapore [20], and Pakistan [21] — each with strong leaders but differing reform durations. Sustained leadership in Türkiye and Singapore embedded lasting change, while Pakistan's shorter reform period saw reversals post-Musharraf (1999–2007), despite significant progress in economic growth, trade, and education Majeed (2014), highlighting the importance of reform longevity for institutional durability.

To assess this empirically, I use an event study around known leadership transitions, providing a quasi-experimental framework to identify their impact on corruption and social trust Chalendard et al. (2023); Kwoka et al. (2016).

Direct measures of societal corruption norms are unavailable over long historical periods—e.g., WVS data do not extend far enough to capture major episodes in Singapore, Türkiye, or Botswana. However, I present evidence consistent with Hypothesis 2 by using proxies and tracing signs of hysteresis in corruption levels and broader social norms. I use broader proxies: public sector corruption and social trust, the latter widely viewed as both a persistence mechanism and cultural equilibrium Giuliano and Nunn (2021).

As such, this evidence does not conclusively prove the hysteresis effect changes social norms on corruption but does provide evidence that is consistent with the hysteresis effect, and it does provide evidence for hysteresis in public sector corruption and social trust. I use four key indicators from the Varieties of Democracy (V-Dem) dataset. To proxy corruption norms, I use two indicators. *Public Sector Theft* (v2exthftps) measures embezzlement by officials, while *Public Sector Corrupt Exchanges* (v2excrptps) captures bribery and favoritism in state dealings. A *Social Trust Proxy*, averaging equal treat-

[20]South Korea, China, and Botswana also exemplify sustained, high-intensity leadership that produced lasting institutional reform—consistent with the hysteresis effect.

[21]By contrast, Afghanistan and Myanmar illustrate further failed reform episodes. In both cases, reforms lacked intensity and duration? Afghanistan due to weak enforcement despite prolonged support, and Myanmar due to democratic backsliding following the 2021 military coup.

ment (v2pepwrsoc) and impartial administration (v2clrspct), reflects societal fairness and institutional trust. Higher values indicate greater public sector integrity and lower levels of corruption [22].

Event Study Regression

To formally test H2, I estimate the following event study regression:

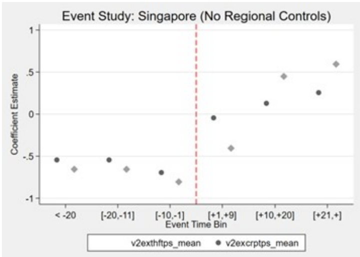
$$Y_t = \beta_0 + \sum_{k \neq 0} \beta_k \cdot [\text{event\_bin}_{t=k}] + \varepsilon_{it} \tag{E6}$$

Let Y_t denote the outcome of interest, where I proxy for corruption in the public sector using the *Public Sector Theft Index*, *Public Sector Corrupt Exchanges Index*, and *Social Trust Proxy*, separately. Event time is grouped into bins relative to the reform year. ε_{it} is the error term.

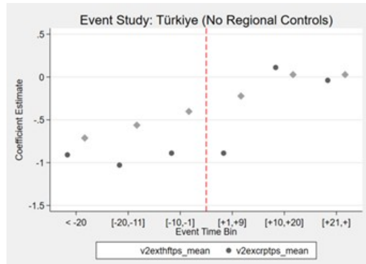
(Figure 6a–6f) presents regression estimates, with the dashed line showing the start of reformist regime. These results suggest that corruption and social trust improvements in Türkiye and Singapore endured well beyond the tenure of Atatürk and Lee Kuan Yew, while Pakistan's gains under Musharraf were short-lived, consistent with Hypothesis 2 [23]. In Pakistan, even though corruption indicators showed some short-run improvements, the social trust proxy remained unchanged—suggesting that the regime was too short-lived to shift underlying societal norms. The contrast highlights the importance of sustained and intensive leadership in generating lasting institutional change.

[22] Both variables are preferred over WGI variables for the longer time periods as WGI variables are not available for earlier periods.

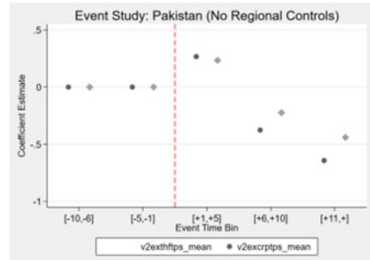
[23] Regression tables available on request.



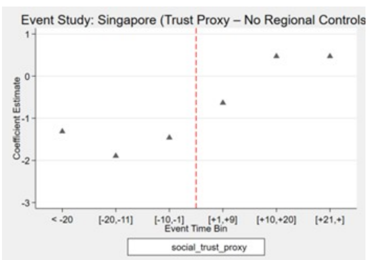
(a) Panel A: Singapore. Corruption Indicators



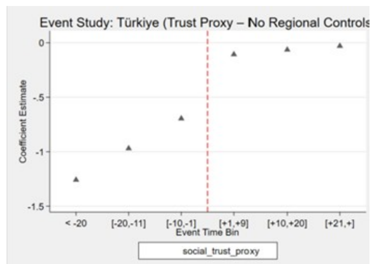
(b) Panel B: Türkiye. Corruption Indicators



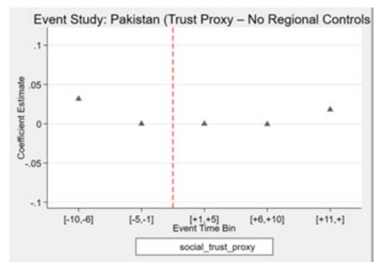
(c) Panel C: Pakistan. Corruption Indicators



(d) Panel D: Singapore. Trust Proxy



(e) Panel E: Türkiye. Trust Proxy



(f) Panel F: Pakistan. Trust Proxy

Figure 6: Figure 6: Event Study Results — Corruption and Trust Indicators by Country

Robustness

To ensure the results reflect country-specific effects rather than regional trends, I control for regional averages (excluding the country's own value), and the findings remain robust. Results available on request.

7 Conclusion

This paper began with a set of simple but important questions: How do institutions evolve? Why are they so persistent? And why are successful institutional transformations so rare, limited to outlier cases such as Singapore, Türkiye, South Korea, Botswana, and China?

While existing models (e.g., Acemoglu et al., 2001; North, 1990) emphasize institutional persistence, they do not explicitly model the dynamic interaction between leadership and population norms. To address this gap, I propose a simple framework to conceptualize how a country's institutions evolve through the interaction between population norms and leadership traits.

The framework is grounded in two new concepts: the Population–Leadership Symmetry Principle and the Leadership Hysteresis Effect. The former explains institutional persistence: leadership traits mirror prevailing societal norms, shaped by incentives, habit formation, selection pools, and normative expectations. The latter shows how this inertia can be overcome—albeit rarely—through sustained, reformist leadership capable of gradually reshaping societal norms. In both concepts, societal norms play a central role, acting as the unifying factor between them.

Drawing on formal modelling and cross-country empirical tests, I document strong alignment between societal norms and leadership integrity, consistent with the symmetry principle. I also provide initial evidence that durable institutional change emerges from leadership episodes that are both intense and long-lasting, in line with the hysteresis mechanism. This framework is also calibrated to notable cases of reform episodes.

By offering a unified framework that accounts for both persistence and change, this paper helps explain why most societies remain trapped in cycles of weak governance, and how a few have broken free of this inertia—with important policy lessons for future reformers.

Although this paper focuses on corruption, the logic of the Population–Leadership Symmetry Principle may apply to other dominant societal values. For instance, it may be the case that in a highly religious society, a religious leader is more likely to emerge; in a society shaped by communist ideals, a communist leader is more probable. These potential applications may suggest that leadership traits tend to reflect a society's most salient values—a possibility that warrants further exploration.

References

Acemoglu, D. and Jackson, M. O. (2015). History, expectations, and leadership in the evolution of social norms. *The Review of Economic Studies*, 82:423–456.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Acemoglu, D. and Robinson, J. A. (2025). Culture, institutions, and social equilibria: A framework. *Journal of Economic Literature*, 63(2):637–692.

Aidt, T. S. (2009). Corruption, institutions, and economic development. *Oxford Review of Economic Policy*, 25(2):271–291.

Ajzenman, N. (2021). The power of example: Corruption spurs corruption. *American Economic Journal: Applied Economics*, 13:230–257.

Andvig, J. C. and Moene, K. O. (1990). How corruption may corrupt. *Journal of Economic Behavior Organization*, 13:63–76.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297.

Ashraf, Q. and Weil, D. N. (2024). *Economic Growth*. Taylor Francis.

Ayuso Castillo, M. (2020). Political change from mao zedong to deng xiaoping. In *Proceedings of the Political Reform Conference*.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Bicchieri, C., Dimant, E., and Xiao, E. (2022). Deviant or wrong? the effects of norm information on the efficacy of punishment. *Journal of Economic Behavior Organization*, 195:270–287.

Borjas, G. J. (2000). The economics of immigration. *Issues in the Economics of Immigration*.

Breunig, R. and Majeed, O. (2020). Inequality, poverty and economic growth. *International Economics*, 161:83–99.

Burguet, R., Ganuza, J.-J., and Garcia Montalvo, J. (2016). The microeconomics of corruption: A review of thirty years of research. *Universitat Pompeu Fabra Working Paper*.

Chalendard, C., Fernandes, A. M., Raballand, G., and Rijkers, B. (2023). Corruption in customs. *The Quarterly Journal of Economics*, 138:575–636.

Cingano, F. (2014). Trends in income inequality and its impact on economic growth. *OECD Social, Employment and Migration Working Papers*, (163).

Fisman, R. and Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy*, 115(6):1020–1048.

Gächter, S. and Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, 531:496–499.

Gao, Y. and Yao, L. (2016). Anti-corruption in the contemporary era: Insights from deng xiaoping's theories. In *2nd International Conference on Humanities and Social Science Research (ICHSSR 2016)*, pages 155–158. Atlantis Press.

Giuliano, P. and Nunn, N. (2021). Understanding cultural persistence and change. *The Review of Economic Studies*, 88:1541–1581.

Gulino, G. and Masera, F. (2023). Contagious dishonesty: Corruption scandals and supermarket theft. *American Economic Journal: Applied Economics*, 15:218–251.

Halter, D., Oechslin, M., and ZweimÃ¼ller, J. (2014). Inequality and growth: The neglected time dimension. *Journal of Economic Growth*, 19:81–104.

Hodge, A., Shankar, S., Rao, D. P., and Duhs, A. (2011). Exploring the links between corruption and growth. *Review of Development Economics*, 15:474–490.

Jain, A. K. (1998). Models of corruption. In *Economics of Corruption*, pages 13–34. Springer.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2005). Governance matters iv: Governance indicators for 1996–2004. *World Bank Policy Research Working Paper*, (3630).

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2010). The worldwide governance indicators: Methodology and analytical issues. *World Bank Policy Research Working Paper*, (5430).

Keizer, K., Lindenberg, S., and Steg, L. (2008). The spreading of disorder. *Science*, 322(5908):1681–1685.

Kuru, A. T. (2009). *Secularism and State Policies Toward Religion: The United States, France, and Turkey*. Cambridge University Press.

Kwoka, J., Hearle, K., and Alepin, P. (2016). From the fringe to the forefront: Low cost carriers and airline price determination. *Review of Industrial Organization*, 48:247–268.

La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2008). The economic consequences of legal origins. *Journal of Economic Literature*, 46(2):285–332.

Liu, H. and Zhang, H. (2017). Lee kuan yewâs thoughts on talent and singaporeâs development strategy. In *Lee Kuan Yew Through the Eyes of Chinese Scholars*.

Majeed, O. (2014). External debt, bankruptcy and pakistan. *Pakistan Today*.

Mauro, P. (2004). The persistence of corruption and slow economic growth. *IMF Staff Papers*, 51(1):1–18.

Morton, F. (2018). Quett ketumile joni masire: Some notes by his biographer. *Botswana Notes and Records*, 50:349–355.

Morton, F. and Ramsay, J. (1994). *The Making of a President: Sir Ketumile Masire's Early Years*. Pula Press Gaborone.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

Nunn, N. (2023). The persistence paradox: Critical junctures and our shared future.

Pamuk, Å. (2007). Economic change in twentieth century turkey: Is the glass more than half full? In *The Cambridge History of Turkey*, volume 4, pages 266–300.

Persson, A., Rothstein, B., and Teorell, J. (2013). Why anticorruption reforms fail: Systemic corruption as a collective action problem. *Governance*, 26(3):449–471.

Reno, W. (1998). *Warlord Politics and African States*. Lynne Rienner Publishers.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system gmm in stata. *The stata journal*, 9(1):86–136.

Sterk, V. (2016). The dark corners of the labor market. *CFM Discussion Paper*.

Trevaskes, S. (2002). Courts on the campaign path in china: Criminal court work in the ”yanda 2001” anti-crime campaign. *Asian Survey*, 42:673–693.

Weisel, O. and Shalvi, S. (2015). The collaborative roots of corruption. *Proceedings of the National Academy of Sciences*, 112:10651–10656.

Wigner, E. P. (1964). Symmetry and conservation laws. *Proceedings of the National Academy of Sciences*, 51(5):956–965.

Wigner, E. P. (1965). Violations of symmetry in physics. In *Philosophical Reflections and Syntheses*, pages 343–358. Springer.

Williams, D. (2014). The leadership of singaporeâs lee kuan yew and the implications for leadership development. Center for Public Leadership, Kennedy School of Government, Harvard University. https://www.globalleadership.net/uploads/the_leadership_of_singapores_lee_kuan_yew__the_implications_for_leadership_development.pdf.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of Econometrics*, 126:25–51.

Appendix Table A1: Country List

Table A1: Country List — Initial Social Values with No Change Over Time vs. Full Sample for Pooled OLS

Initial Norms Sample	**Initial Norms Sample**	**Full Sample for Pooled OLS**	**Full Sample for Pooled OLS**	**Full Sample for Pooled OLS**
Andorra	Trinidad and Tobago	Algeria	Japan	Spain
Argentina	Turkiye	Andorra	Jordan	Sweden
Australia	Ukraine	Argentina	Kazakhstan	Switzerland
Brazil	United Kingdom	Armenia	Kenya	Tajikistan
Canada	United States	Australia	Korea, Rep.	Thailand
Chile	Uruguay	Azerbaijan	Kuwait	Trinidad and Tobago
China	Vietnam	Bangladesh	Kyrgyz Republic	Tunisia
Colombia		Belarus	Lebanon	Turkiye
Cyprus		Bolivia	Libya	Ukraine
Egypt, Arab Rep.		Brazil	Macao SAR, China	United Kingdom
Ethiopia		Bulgaria	Malaysia	United States
Georgia		Burkina Faso	Maldives	Uruguay
Germany		Canada	Mali	Uzbekistan
Ghana		Chile	Mexico	Venezuela, RB
Guatemala		China	Moldova	Vietnam
Hong Kong SAR, China		Colombia	Mongolia	West Bank and Gaza
India		Cyprus	Morocco	Yemen, Rep.
Indonesia		Czech Republic	Myanmar	Zambia
Iran, Islamic Rep.		Ecuador	Netherlands	Zimbabwe
Japan		Egypt, Arab Rep.	New Zealand	
Jordan		Estonia	Nicaragua	
Korea, Rep.		Ethiopia	Nigeria	
Malaysia		Finland	Norway	
Mexico		France	Pakistan	
Morocco		Georgia	Peru	
Netherlands		Germany	Philippines	
New Zealand		Ghana	Poland	
Poland		Greece	Puerto Rico	
Romania		Guatemala	Qatar	
Russian Federation		Haiti	Romania	
Rwanda		Hong Kong SAR, China	Russian Federation	
Serbia		Hungary	Rwanda	
Slovenia		India	Serbia	
South Africa		Indonesia	Singapore	
Spain		Iran, Islamic Rep.	Slovak Republic	
Sweden		Iraq	Slovenia	
Thailand		Italy	South Africa	

Table A2: Determinants of Corrupt Leadership — OLS

VARIABLES	**(1)** Rule of Law	**(2)** Rule of Law	**(3)** Rule of Law	**(4)** Rule of Law	**(5)** Rule of Law	**(6)** Rule of Law
Societal values: Illegal benefits	**-0.37*** **	**-0.26*** **				
	(0.08)	(0.06)				
Societal values: Cheating on taxes			**-0.36*** **	**-0.24*** **		
			(0.10)	(0.08)		
Societal values: Fare evasion					**-0.34*** **	**-0.27*** **
					(0.09)	(0.06)
Education expenditure (% GDP)		0.08*		0.09**		0.10**
		(0.04)		(0.04)		(0.04)
Trade openness (% GDP)		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)
Log of GDP per capita		0.72***		0.74***		0.76***
		(0.07)		(0.07)		(0.07)
Tax revenue (% GDP)		0.01		0.02*		0.01
		(0.01)		(0.01)		(0.01)
Constant	1.36***	-6.55***	1.23***	-7.08***	1.30***	-7.02***
	(0.24)	(0.72)	(0.27)	(0.70)	(0.26)	(0.65)
Observations	178	115	176	114	175	113
R-squared	0.14	0.71	0.10	0.70	0.11	0.72
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls		Yes		Yes		Yes

Note: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A3: Determinants of Corrupt Leadership — GMM — Democracy and Freedom of Expression

VARIABLES	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
	Con. Crpt	Con. Crpt	Con. Crpt	Con. Crpt	Con. Crpt	Con. Crpt
Initial societal values: Claiming illegal benefits	**-0.55*****	**-0.33****				
	(0.17)	(0.13)				
Societal values: Claiming illegal benefits = D,		**0.06**				
		(0.08)				
Initial societal values: Fare evasion			**-0.45*****	**-0.50*****		
			(0.17)	(0.14)		
Societal values: Fare evasion = D,				**0.10**		
				(0.20)		
Initial societal values: Cheating on taxes					**-0.35****	**-0.45****
					(0.14)	(0.18)
Societal values: Cheating on taxes = D,						**0.40**
						(0.35)
Education expenditure (% GDP)		0.04		0.28		0.19
		(0.15)		(0.17)		(0.16)
Trade openness (% GDP)		0.01**		0.01		0.01
		(0.00)		(0.00)		(0.00)
Log of GDP per capita		0.51*		0.38		0.34
		(0.28)		(0.29)		(0.40)
Electoral democracy index		3.30		2.19		1.26
		(2.97)		(3.35)		(3.21)
Freedom of Expression and Alternative Sources of Information index		-2.34		-1.41		-0.39
		(2.77)		(2.72)		(2.55)
Constant	**1.82*****	-4.70	**1.59*****	-4.02	**1.18*****	-3.68
	(0.49)	(2.84)	(0.52)	(2.84)	(0.40)	(3.86)
Observations	220	67	216	63	220	65
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Countries	55	37	54	36	55	36

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Determinants of Corrupt Leadership — GMM with V-Dem (Public Sector Theft)

VARIABLES	**(1)** PS Theft	**(2)** PS Theft	**(3)** PS Theft	**(4)** PS Theft	**(5)** PS Theft	**(6)** PS Theft
Initial societal values: Claiming illegal benefits	**-0.45***** (0.14)	**-0.24*** (0.13)				
Societal values: Claiming illegal benefits = D,		**0.01** (0.06)				
Initial societal values: Fare evasion			**-0.38***** (0.14)	**-0.35**** (0.14)		
Societal values: Fare Evasion = D,				**-0.06** (0.18)		
Initial societal values: Cheating on Taxes					**-0.34**** (0.14)	**-0.33**** (0.14)
Societal values: Cheating on taxes = D,						**0.06** (0.26)
Trade Openness (% GDP)		0.00 (0.00)		0.00 (0.00)		0.00 (0.00)
Log of GDP per capita		0.78*** (0.21)		0.60*** (0.16)		0.60*** (0.18)
Constant	**3.84***** (0.39)	-4.56** (2.16)	**3.70***** (0.41)	-3.11* (1.55)	**3.46***** (0.36)	-2.93* (1.71)
Observations	215	67	211	63	215	65
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of country_id	54	37	53	36	54	36

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.