Australian National University
Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Why Institutions Endure: Norms, Leadership, and What Enables Reform

CAMA Working Paper 36/2025
June 2025 (Revised in August 2025)

Omer Majeed
Centre for Applied Macroeconomic Analysis, ANU

Abstract

Why do trapped countries remain locked in bad institutions while a few escape? This paper introduces a novel framework in which societal norms and leadership traits drive institutional persistence and change. It offers four insights. First, leadership traits correlate with societal norms through contagion, incentives, selection pool effects, and normative alignment — cross-country tests are consistent with this framework. Second, it explains why sustainable reforms are rare. Third, it shows reforms endure only when norms shift and explains the mechanism. Fourth, in notable transformations — Singapore, South Korea, Botswana, and Turkiye — it shows enduring reforms require both duration and intensity.

Keywords

Institutions, societal norms, development, leadership, economic growth

JEL Classification

D02, D73, D72, E02, H83, O43, O57, Z10, Z18

Address for correspondence:

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

ISSN 2206-0332

The Centre for Applied Macroeconomic Analysis in the Crawford School of Public Policy has been established to build strong links between professional macroeconomists. It provides a forum for quality macroeconomic research and discussion of policy issues between academia, government and the private sector.

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

Why Institutions Endure: Norms, Leadership, and What Enables Reform *

Omer Majeed
Centre for Applied Macroeconomic Analysis, ANU

May 2025

Abstract

Why do trapped countries remain locked in bad institutions while a few escape? This paper introduces a novel framework in which societal norms and leadership traits drive institutional persistence and change. It offers four insights. First, leadership traits correlate with societal norms through contagion, incentives, selection pool effects, and normative alignment — cross-country tests are consistent with this framework. Second, it explains why sustainable reforms are rare. Third, it shows reforms endure only when norms shift and explains the mechanism. Fourth, in notable transformations — Singapore, South Korea, Botswana, and Turkiye — it shows enduring reforms require both duration and intensity.

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

JEL Codes: *D02, D73, D72, E02, H83, O43, O57, Z10, Z18*

*I would like to thank Flavio Menezes, Petr Sedlacek, John Asker, Gabriele Gratton, Günther Schulze, Gabrielle Watson, Jenny Gordon, Warwick McKibbin, Jane Golley, Hal Hill, Prema-chandra Athukorala, 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

1 Introduction

Institutions are one of the most important drivers of a society's level of development. But what defines institutions, why do they persist, and when can they change? Most countries trapped in low levels of development remain there. Corrupt and extractive institutions are the primary barrier to sustainable economic development Acemoglu et al. (2001); Mauro (2004); Ashraf and Weil (2024); Kaufmann et al. (2010, 2005). Yet a few countries, Singapore, South Korea, China, Botswana, and Türkiye have achieved dramatic institutional transformation. So what explains their transformation? Acemoglu et al. (2001) highlight institutional persistence, especially the enduring effects of colonial rule. The above country examples then present a puzzle: they underwent deep, lasting institutional reform despite not fitting neatly into the colonial origin story (Acemoglu et al., 2001). As such, while colonial rule may explain a lot of country level variation, it may not give the full picture. This paper studies the drivers of institutions, and draws policy lessons from the notable cases of transformation.

This paper develops a novel model showing how the interplay between societal norms and leadership traits shape institutions [1]. It complements recent work that identifies norms as central to economic development (Giuliano and Nunn, 2021) and institutions (Acemoglu and Robinson, 2025), and that culture plays a decisive role in how democracies operate (Gratton et al., 2025), while addressing a critical gap in the literature. Acemoglu and Robinson (2025) argue norms define institutions but note existing models lack a formal mechanism incorporating norms. This paper offers a more formal model, illustrating how norms shape institutions. The paper starts by showing how a principal–agent (PA) framework can be adapted to give the microfoundations of institutions, illustrating how corruption levels modify incentives—via accountability probabilities and moral costs—to drive outcomes. The framework examines why institutional change is rare and, when it does occur, employs a difference equation framework to trace the evolution of norms and the role of policy in steering them. While prior frameworks such as Acemoglu et al. (2001) and North (1990) emphasize persistence, they do not model the dynamic feedback between leadership and social norms, and thus overlook their co–evolution.

The paper hypothesizes that leadership traits correlate with societal norms, as social behavioural pressures influence individuals, including leaders—via contagion effects, habit formation, incentives, selection pool effects, and normative alignment. Together, these mechanisms lead to a substantive new conclusion: a behavioral symmetry between the norms of the population and leaders, which this paper labels the *Population—Leadership Symmetry Principle*. Consequently, reform is rare: in societies skewed toward corrupt norms, reformist leadership emergence is unlikely. In the long run, leadership emergence is modeled as a stochastic process conditioned on the so-

[1] The terms social and societal are used interchangeably in this paper

cietal distribution of norms. The paper hypothesizes that when reformist leaders do arise, they can only achieve lasting reform if they are able to shift the underlying norms through intensity and duration of their efforts, a process this paper labels *Leadership Hysteresis Effect*. The paper then demonstrates how such transformations unfold and identifies the tools used in major reforms.

Importantly, viewing institutions with this framework has another advantage: it can be used to study how bad leaders, on rare occasions, can exploit reforms to erode norms, and it shows that these leaders are most dangerous when they can combine regime longevity with reform intensity.

Empirically, the paper uses panel data and a quasi-experimental approach, producing results consistent with the *Population–Leadership Symmetry Principle* and the *Leadership Hysteresis Effect*. It links societal norms to macro-level corruption outcomes—something that has been lacking in the literature (Acemoglu and Robinson, 2025). The paper shows how weak social norms, constructed from individual level data, are correlated with corrupt leadership in government. Finally, the model is calibrated to notable cases of institutional transformation: Singapore under Lee Kuan Yew, Botswana under Sir Seretse Khama, Türkiye under Mustafa Kemal Atatürk, and South Korea from the early 1990s to the late 2000s. This demonstrates its applicability to diverse, real world examples. The calibration exercise also contributes to the debate on how long norms take to change. While Acemoglu and Robinson (2025) suggest a shift can occur within a generation, Nunn (2023) argues it may take several. The results here indicate that for corruption–related norms, sustainable change can occur over 20–25 years, aligning more closely with Acemoglu and Robinson (2025) than Nunn (2023).

The framework of this 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 governing institutional outcomes and policy—filling an important gap in the literature (Acemoglu and Robinson, 2025).[2]

The paper proceeds by examining microdata evidence. The hypothesis that societal norms on corruption shape the attitudes and behaviors of people is bolstered by micro-level evidence. For instance, Gulino and Masera (2023) show that news of local corruption scandals significantly increases the probability of theft at supermarkets, showcasing contagion effects of societal norms. Similarly, Ajzenman (2021) show that revelations of corruption by local officials lead to increased cheating by secondary schools students taking tests. Evidence also suggests that people in positions of power emulate the characteristics of their country. Fisman and Miguel (2007) find that diplomats from more corrupt countries had significantly more unpaid New York City parking violations (which fell sharply once immunity was lifted) illustrating habit formation and incentive responsiveness in corrupt behavior.

[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.'

Similarly, Gächter and Schulz (2016) show through cross-sectional evidence from 23 countries that the prevalence of rule violations in a society impairs 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); Borjas (2000) show that societal values are an important determinant of norm compliance.

This paper builds on this micro-level evidence by emphasizing that, at the macro level, societal norms exert pressure and can significantly shape the traits of individuals, including 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 PA framework, prevalent corruption weakens enforcement and incentivizes corruption. High-corruption environments overwhelm enforcement systems, lowering capacity and reducing the likelihood any individual case will be punished. 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). Thus corruption is fostered in corrupt societies[3].

Third, the prevalence of integrity in the general population influences leadership selection: in societies with a larger share of non–corrupt individuals, future leaders, policymakers and judges are more likely drawn from honest segments, a process referred to here as the selection pool effect. Fourth, societal preferences exert normative pressure on leadership. Populations which strongly favor honesty and good governance, demand greater integrity in leadership. Over time, when societal preferences and leadership behavior become aligned, a stable equilibrium emerges.

Together, these mechanisms lead to an important conclusion: leadership traits are correlated with those of the population. The paper shows how societal norms and leadership traits can be used to define a society's steady state of institutions. It shows why meaningful institutional change is rare as it relies on a shift in underlying norms.

Given this, how do leaders in rare cases implement successful reforms? Drawing on notable cases, the paper shows when and how reforms come about, and uses a difference-equation framework to show how targeted policies alter a society's corruption distribution. It demonstrates the importance of reform intensity and duration in the hysteresis effect. The paper explains why such episodes are exceptionally rare.

[3]Various factor can be disruptive —such as wars, international sanctions, regional turmoil, etc. However, the paper argues that even with disrupting factors, societies where corrupt norms are deeply entrenched are less likely to see current or future governing institutions free themselves from corruption.

This framework aligns with North (1990), who argues that institutions evolve slowly and that 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 a closely related view to this paper, showing how rare, prominent leaders can 'counteract history' and shift social expectations, where they use Nelson Mandela as an example. Similarly, Guiso et al. (2016) show that civic capital rises with the duration and intensity of medieval self-governance.

This research complements the cultural–institutional perspective of Acemoglu and Robinson (2025), whose notable contribution is to argue that social norms shape institutional equilibria. They examine cultural attitudes through a 'systems approach,' conceptualizing culture as a set of attributes, the distribution of those attributes, and their configuration. They describe this framework as a 'preliminary' step toward understanding how culture can be part of institutional equilibrium, and explicitly call for more formal modeling. This paper offers a more formal model, showing how norms shape institutions [4].

The paper starts by showing how the PA framework can be used to illustrate the microfoundations of institutions and the incentives at play in society. PA models often analyze corruption through incentives like detection probability and rewards (Jain, 1998), 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).

The empirical analysis offers evidence consistent with the symmetry principle and hysteresis effect, with a focus on the symmetry principle. It uses World Values Survey (WVS) person level data to construct societal integrity measures (e.g., attitudes toward tax evasion, fare dodging, benefit fraud). For leadership corruption, the paper uses the Worldwide Governance Indicators (WGI) control of corruption index, which captures both petty and grand corruption in government and leadership. Findings remain robust to various methods (OLS, system GMM, using log lags for reverse causality), and alternative data sources and variables. A key limitation is the lack of a valid external instrument, and as Acemoglu and Robinson (2025) cautions that causality is hard to establish in this field, results are interpreted as associations, not causal effects.

The paper exploits quasi–experimental variation in leadership transitions to contrast Singapore (under Lee Kuan Yee) and Türkiye's (under Kemal Ataturk) sustained reformist tenures with Musharraf's brief Pakistan rule, providing evidence consistent with the hysteresis effect.

[4]The model examines the distribution of corrupt norms held by the population and its impact on leadership traits and institutions. Using contagion effects, habit formation, incentives, and behavioral traits, the paper shows how norms shape institutional equilibrium, as association Acemoglu and Robinson (2025) argue is missing in the literature.

The paper proceeds as follows: Section 2 outlines key concepts; Section 3 provides micro–foundations (PA framework); Section 4 presents the main model; Section 5 calibrates reform episodes; Section 6 presents the data and results; Section 7 concludes.

2 Model set up and definitions

This paper argues that institutions are run by people, and that the norms of those who run them influence how institutions function. As such, institutions reflect the norms of the society they are embedded in.

This section presents key concepts and definitions used to clarify how norms matter, and how institutional equilibria emerge, persist, and, in some cases, evolve. The paper starts by defining the core variables and parameters. The paper then, with the help of a PA framework, lays down the micro–foundations, showing how incentives vary for individuals — including leaders — as corruption levels change within society.

The paper then develops the core mechanisms that show how norms define equilibria and how leadership traits mirror societal norms. These mechanisms include incentives, contagion effects, selection effects, and normative alliance. The paper then uses a difference equation framework to show how, in rare cases, leaders alter societal norms and shift the equilibrium of society.

Corruption

Following Gulino and Masera (2023), the paper defines 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.

As such, when there is corruption in leadership, leaders are concerned more with their private gains than economic development to improve welfare of others in society.

Population characteristics θ_t^{NC}

The paper assumes 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. An alternative though slightly less important 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 of all layers of leadership, including top leadership, plus other tiers such as judges, policy makers, politicians etc[5] . 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 thus more focused on the development of growth-enhancing institutions. While $C_A = \underline{A}$ represents the opposite: corruption, self–interest and institutional decay. In part, $C_{A,t}$ is seen by the public as a signal of what is acceptable. 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

The framework considers 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}). A motivation for this formulation is that institutions do not exist in isolation: they reflect the norms of the people who occupy them.

Institutions can be conceptualized as complimentary inputs of two main factors (Fortunato and Panizza, 2015, Krieger, 2022). The framework assumes a Cobb-Douglas [6] 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

[5]With top leadership having more weight.

[6]Institutions production function can also be made linear as per Krieger (2022) without loss of generality.

drives institutional change. Ψ could reflect leadership effectiveness, including internal harmony among layers of leadership and external factors [7] [8] .

3 Micro-foundations — Principal-Agent Model

This section provides a micro-foundation for how societal corruption incentivizes corrupt behavior, including of 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.

The model argues that both leadership and citizens take social cues, i.e. enforcement signals (P_D) and internalized moral costs (m_i) from the wider society. While the model is framed around agents such as politicians, policy officials, or bureaucrats in leadership roles, the underlying incentive structure applies equally to the broader population.

To explore the incentive dynamics, the PA framework is applied, highlighting how leadership, judicial enforcement (J) [9], contagion effects, and corruption incentives can impact outcomes as norms vary in the society. 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

This section shows how the PA framework can be adapted to give the micro foundations of institutions. The paper extends Mauro (2004)'s two-period model to show how widespread corruption fosters inertia. It illustrates that by endogenizing variables like the probability of detection and moral costs to societal norms, the PA framework can capture how society shapes individual incentives, thereby imposing mean reversions and explaining institutional inertia. It also clarifies what is required to counteract this inertia, while the section on the probability of change explains why counteraction is

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

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

[9] There are several ways to model J. $J_{t+1} = aJ_t + \beta C^M_{A,t} - \gamma\theta^C_t$, suggests judicial inertia can be overcome by strong leadership effort, but a corrupt population will weaken judicial reforms. However, for the purposes of this paper, this section does not need to formalise J.

likely to be rare. Theory of change in norms is developed further in the section on the role of leadership.

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

Model Setup

- Principal: Population
- Agents: $C^M_{A,t}$, $C_{A,t}$
- 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: $P_D = f(C_{A,t}, \theta^{NC}_t)$ (see below)
- Penalty if caught: Must return all bribes received and pay an additional fine $\widetilde{T}$

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. Similar to Mauro (2004) for simplicity, we have isoelastic utility, and there is no discounting.

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:

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

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 and 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.

Importantly, moral costs (m_i) can vary in a society based on corruption for two reasons. First, 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). The underlying notion is that corruption in society lowers the moral guilt associated with corrupt behavior.

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 changes in P_D and m_i reflect 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:

$$P_D, m_i = f(C^M_{A,t}, \theta^{NC}_t) \tag{5}$$

In normal times, when reformist leadership is absent, θ^{NC}_t influence both P_D and m_i. When $C^M_{A,t}$ and θ^{NC}_t are low, low detection rates embolden corrupt behavior, entrenching bad institutional inertia. The reverse is true for strong leadership and honest countries.

However, under strong reformist leadership (where $C^M_{A,t} > \theta^{NC}_t$), the framework assumes that leadership 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 Lee Kuan Yew in Singapore exemplify this mechanism Quah (2022).

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

As such, the framework assumes the probability of detection can take different values:

$P_D = P_l$ (low detection probability). In a corrupt society, with weak governance, $P_l = \theta^{NC}_t$, when θ^{NC}_t is low. Similarly, m_i is lower when θ^{NC}_t 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^{NC}_t is high, or reformist tendencies (high $C^M_{A,t}$) 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 generate 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 generate a low corruption equilibrium, where $C_{A,t}$ is incentivized to be behave honestly

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

Implications for the society

These results imply that macro–level prevalence of corruption has implications at the individual (micro) level, as corruption in society distorts incentives by altering the probability of accountability and the moral costs faced by individuals, including leaders, in the society. Corruption then perpetuates corrupt behavior. Conversely, when non-corrupt individuals are more prevalent, honesty is reinforced.

4 Framework

Leadership–Population Symmetry Principle and Steady State

The framework emphasizes that the presence of a symmetry between leadership and population characteristics, with 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 any individual case will be prosecuted. Tit-for-tat tolerance among corrupt actors further undermines deterrence, as mutual complicity discourages reporting and punishment and even extends to rationalizing corrupt behavior as a defensive response to avoid exclusion. 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.

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 — which this paper labels 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 those 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. At the core of the symmetry principle is the idea that societal norms embody incentive structures, contagion effects, selection pool effects, and normative alignment, all of which influence individuals, including those who rise to leadership. As such, traits of the leaders' mirror those of the society. This dynamic gives rise to a self-reinforcing institutional equilibrium. Importantly, mechanisms such as contagion effects, habit formation, and incentives also help explain institutional persistence.

Societal Norms and Leadership Alignment

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

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

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 also more likely to be corrupt.

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

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 non-corrupt to total 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 occurs when θ_t^{NC} and $C_{A,t}$ intersect at the 45-degree line, defining the steady state.

In steady state we have:

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

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. This framework can support multiple steady states, akin to Sterk (2016). [10]

[10]While the core model focuses on the average level of norms in society, the distribution of norms may also matter. For instance, if polarization increases in a society, a dispersed bi-modal distribution may appear and which in turn will increase variance in the type of leader selected. For instance, a bi-modal (e.g., polarized society) distribution could lead to oscillating leadership types. The U.S. provides a possible example, where both left- and right-leaning leaders are regularly elected. The greater the separation between the modes, the more divergent the leadership outcomes may become, as electoral outcomes are increasingly shaped by the central tendencies of distinct ideological clusters. Figure 1 can be adapted to show how greater polarization leads to more divergent leader selection. Importantly, this variance in leadership is likely to also create investment uncertainty. This framework also illustrates how the model—and the concept of population norms—could be extended to explain leadership outcomes

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.

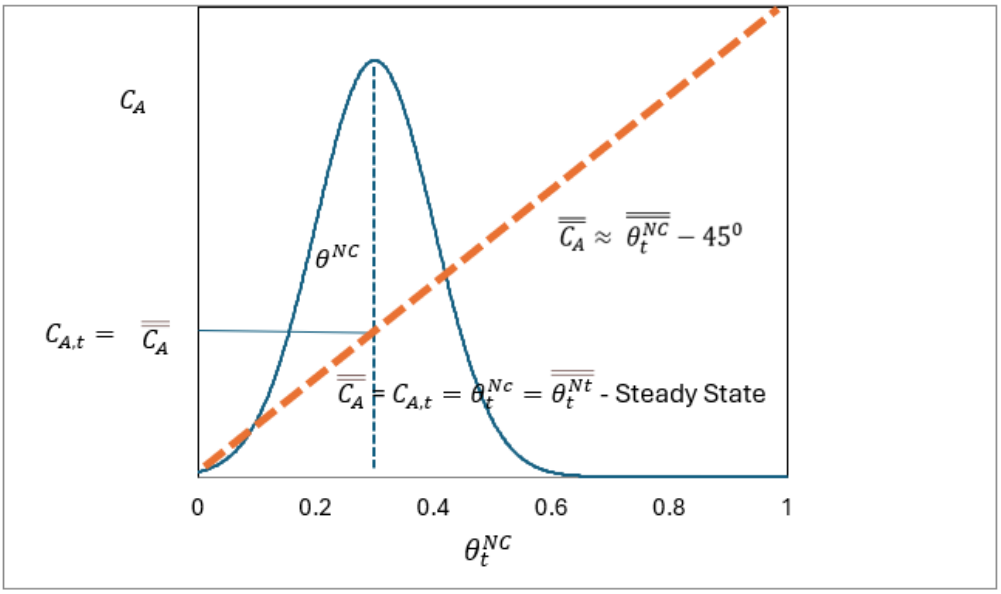


Figure 1: Leadership Population Symmetry Principle

[Result 1: Leadership–Population Symmetry Principle] A behavioral symmetry exists between population norms and leadership norms.

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 national leadership will be correlated with these corrupt social norms.

The Role of Leadership Characteristics, C_A

Given the strong influence of societal norms—operating through incentives, contagion, habit formation, selection pool effects, and normative alignment—meaningful institutional transitions require a shift in norms from corrupt to non-corrupt. Merely

under conditions of polarization driven by rising inequality, social media, or other societal fragmenting forces. While this is an important extension, exploring the distributional aspects requires a paper in itself. Moreover, examining the distribution is not required for the core insights of this paper and is left to future research.

changing laws or formal rules is insufficient, because underlying factors like contagion effects, habit formation, incentives and selection pool effects, are going to exert behavioral pressures and therefore inertia.

This is in line with the arguments of (North, 1990) that institutions evolve slowly. Further, North (1990) points to informal rules and norms often exerting greater influence than formal structures. As such transformation is necessarily slow.

So given this force for inertia, how do institutions change? This section argues that, in rare cases, strong and persistent leadership can overcome inertia and gradually reshape the underlying population characteristics. Further, it shows why this rarely occurs.

In this model, the emergence of leadership is treated as a stochastic process in the long run, but conditioned on the underlying distribution of norms. In a corrupt society, where the population is centered 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, overtime deviations from the mean will emerge and while less like likely there will be instances of reformist leaders coming to power.

These deviations allow for the emergence of reformist leaders (eg., Lee Kuan Yew in Singapore, Deng Xiaoping in China, Sir Seretse Khama in Botswana and Atatürk in Türkiye) who break from the prevailing societal equilibrium. Disruptive leaders may emerge from within the existing leadership structure, as did Deng Xiaoping and successive South Korean leaders, or be imposed through revolution, as with Atatürk.

In these instances, the leadership changed the composition of population norms through education, strong judiciary, strong anti-corruption measures including signaling and legislative reforms, and focus 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 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. Relatedly, Guiso et al. (2016) provide compelling city level evidence on the critical roles of duration and intensity: former medieval free cities in Italy display higher civic capital today, with the effect's magnitude scaling up with the length and strength of past autonomy.

The appearance of a reformist leader (i.e., $C_{A,t} > \theta_t^{NC}$) represents a shock to the system. The model assumes 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 such a meaningful transformation to population characteristics, requires stability and consistency (given by duration) in leadership and reforms. When leadership endures, a *Leadership Hysteresis Effect* can emerge—embedding long-run change that persists beyond the leader's tenure.

The framework argues the following two conditions are necessary and sufficient for institutional change:

1. Sustainable institutional change requires changing the norms 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.

Leadership as a Driver of Norm Change: An Iterative Process

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 respective exit rates 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 individuals removed from 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. During long spells of reformist leadership, natural attrition linked to δ_C becomes increasingly significant, as individuals with entrenched corrupt habits and preferences die out [11].

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}). In highly corrupt societies, corrupt officials may work to force out non-corrupt actors by increasing δ_{NC}.

Singapore, Türkiye, South Korea, China, and Botswana exemplify the key role of ω, μ, δ_C, δ_{NC} in their transitions. In each case, sustained leadership not only enforced legal and institutional reform but also reshaped societal norms through a multipronged 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}. Together, these case studies—and the parameters ω, μ, δ_C, and δ_{NC}—underscore the levers through which sustained leadership can reshape societal norms.

For simplicity, I assume that high $C_{A,t}$ gives $\mu = \sigma$ 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

[11] Even if corrupt leaders and officials are exiled during reformist tenure they can undermine the government from abroad and may return to power b1y leveraging pre-existing networks to re-establish influence. Sustained attrition can gradually weaken this threat

strong leadership increases the proportion of non-corrupt individuals, and vice versa.

Transitional dynamics

In the long run, the system reaches equilibrium when $C_{A,t} = \theta_t^{NC}$. At the steady state, the per cent rates at which corrupt (and non-corrupt) individuals exit and enter the system are the same. Under these conditions, θ_t^{NC} becomes close to 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 $C_{A,t} > \theta_t^{NC}$, increase in the share of non-corrupt population requires:

- 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 variables outnumber constraints, this section imposes additional constraints: $0 \leq \theta_t^{NC} \leq 1$, 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 from the leadership will start to shift the underlying population to the right. That is, if $C_{A,t}$ is large enough for long enough, it will over time shift the underlying population characteristics ($\theta^{NC'}$ to $\theta^{NC''}$).

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.

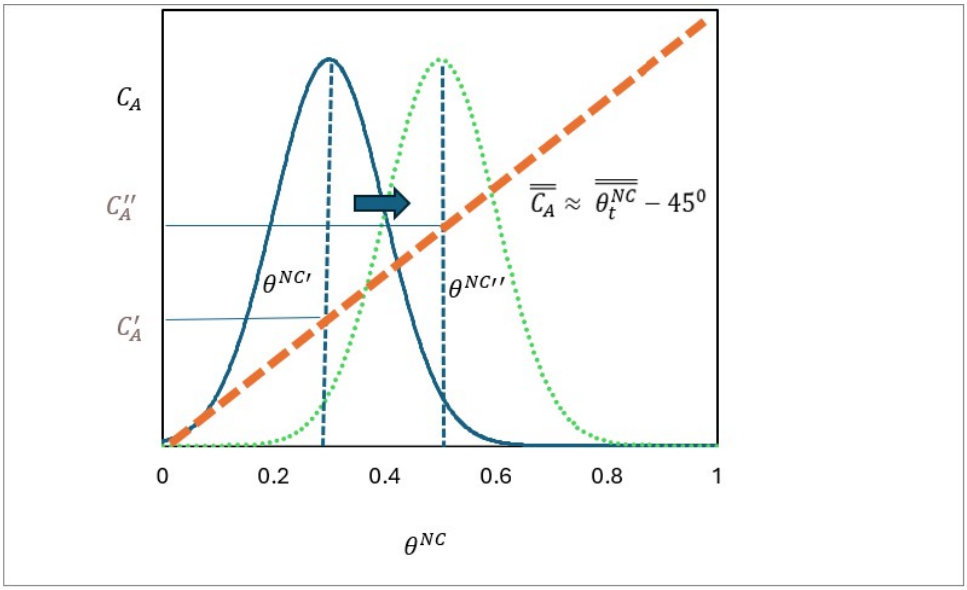


Figure 2: Transition dynamics as leadership reshapes population traits

This again highlights that a meaningful transformation requires the underlying norms on corruption of the population to shift from corrupt to non-corrupt. Key variables like ω, μ, δ_C, δ_{NC}, and therefore norms, are likely to exhibit inertia. Hence, even when formal power changes, informal institutions may take time to follow.

[Result 2: Leadership Hysteresis Effect] Reforms fail when they fail to shift norms. When leadership endures and succeeds in changing population norms, a hysteresis effect emerges, creating durable reforms. Therefore, durable institutional change requires both reform intensity and regime stability—that is sufficient time in office to embed altered social norms.

This result has important policy implications, placing norms at the heart of any sustainable reform initiatives. [12] It also generates several testable hypothesis. One key hypothesis is:

Hypothesis 2: Reforms need intensity and duration to have a lasting effect: only then can they lead to a hysteresis effect. Both intensity and duration are necessary to change underlying norms.

While Acemoglu and Robinson (2025) emphasize the possibility of punctuated shifts in institutional equilibria following leadership or political shocks, Figure 2 illustrates

[12] A small note on longevity and democracies: While one might argue that only autocratic regimes offer the stability and duration required for reformist leaders to shift societal norms, democracies can also exhibit long leadership spells. Leaders such as Nehru (India), Ben-Gurion (Israel), Trudeau (Canada), Merkel (Germany) and remained in office for extended periods, often with successors who maintained broad continuity in governance.

that even after a discontinuous change in leadership of the type $C'_A \rightarrow C''_A$, population norms on corruption will adjust gradually as the mean of the distribution of norms moves right (from $\theta^{NC'}$ to $\theta^{NC''}$), and as society transitions along the 45 degree line. This reflects a hysteresis dynamic, where reform-minded leaders need both norms and the system to move to a new equilibrium. Notably, as the calibration exercise of this paper shows, this can happen within 20-25 years, relatively faster than suggested in previous research Nunn (2023). Where the speed of convergence depends on structural gaps given by $(\mu - \omega)$, $(\delta_{NC} - \delta_C)$, and the starting point of social norms, $\theta^{NC'}$. [13]

When are bad leaders most dangerous?

Importantly, viewing institutions in this framework has an additional advantage: it can highlight when institutions are at risk from bad leaders. The framework suggests bad leaders are most dangerous when they combine tenure duration with reform intensity, thereby increasing the probability of norm erosion.

Like effective reformers, bad leaders can deploy the same tools in the opposite direction to shift norms leftwards in Figure 2. Tools can include: signaling acceptability through official actions, altering educational institutions, shaping media narratives, making strategic legal appointments, promoting like–minded officials, politicizing law enforcement, normalizing fringe beliefs, and attacking independent media and oversight bodies. Viewed this way, a good equilibrium is not necessarily a resilient one. This paper also argues that backsliding is less likely when the conditions for hysteresis — sufficient duration in office and reform intensity are not met or when a threshold is crossed, i.e., when a sufficiently large core of citizens can actively resist norm erosion. This is further explained more in the 'Threshold Effect' subsection.

Probability of Successful Transformation is Low

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 and the consequent probability of reform intensity (P_i) is low. Finally, for 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 in the face of resistance from entrenched elites seek-

[13] 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.

ing 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:

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

The probability of successful reform is low, as it depends on the joint occurrence of three individually unlikely conditions. Given that each component is rare, their product makes success even less likely.

While China, South Korea, Singapore, Türkiye, and Botswana are notable examples of successful institutional reform, they remain 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.

As such, the symmetry principle explains the prevailing (or invariant) tendency—why most societies remain stuck—while the hysteresis effect accounts for rare shifts. As the model formalizes the joint occurrence of P_{NC}, P_i and P_{ld} is uncommon, rendering durable reform episodes as exceptions rather than the rule.

[Result 3: Low probability of change] Durable institutional transformation is unlikely because it requires the rare convergence of non-corrupt leadership, reform intensity, and political longevity. Importantly though, the probability of successful change is greater than zero. [14]

Causality

Leadership typically reflects prevailing societal norms. The symmetry principle implies a directional relationship from societal norms to leadership traits. Result 3 highlights the conditions under which change can occur and explains why it is rare, demonstrating that causality is more likely to run from societal norms to leadership characteristics, thereby reinforcing persistence. Yet in rare instances, leadership can reshape those very norms. The Leadership Hysteresis Effect outlines the mechanisms through which sustained and intensive leadership can drive such transformation.

As such, both are fundamental: the symmetry principle explains persistence in most societies, while the hysteresis effect accounts for the rare cases where norms are successfully transformed.

[14] As shown by cases of notable transformation in the society

Threshold Effect

The paper also assumes a threshold effect: once θ_t^{NC} exceeds a critical value $\hat{\theta}_t^{NC}$, $C_{A,t}$ will have difficulty reversing institutional quality. 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 backsliding and corruption in leadership, applying public pressure—such as protests—to uphold institutional integrity. [15]

Policy implications

By modeling the interplay between societal norms and leadership traits—and identifying the conditions under which reform can succeed—the framework can offer a diagnostic lens for policymakers and international institutions. It shows that durable institutional change requires shifting social norms; legislative reform alone is insufficient. The experiences of countries such as South Korea, Singapore, Botswana, and Türkiye illustrate the possibility of sustained reform. While the paper does not catalogue their specific policy choices in detail, it draws on these cases to highlight mechanisms that others might adopt.

The difference equation framework (along with the calibration exercise of the paper), incorporating variables such as corruption norms, can help estimate the duration and intensity of reform required to shift equilibrium outcomes. This underscores the importance of aligning reform strategies with the underlying distribution of social norms. Finally, the paper shows that a good equilibrium is not necessarily a resilient one, and identifies when and how institutions are most vulnerable to backsliding under adverse leadership. Focusing on positive reforms, these mechanism and pathways are further highlighted in the calibration section.

5 Calibration to Reform-Era Transitions

This section calibrates the model to the average reform trajectory observed in four countries—Singapore, South Korea, Türkiye, and Botswana—during periods of sustained, high-intensity anti-corruption efforts.

These diverse and notable cases illustrate how the model can be applied to real-world settings. The calibration also underscores the model's emphasis on both the intensity and duration of reform, highlighting their joint importance for successful institutional transformation.

[15] An example of this is South Korea in 2024, where President Yoon attempted to impose martial law, but it resulted in mass protests by the people and his ultimate removal from office.

Direct measures of societal corruption norms are unavailable over long historical periods. As such the paper will use the average of public sector corruption and public sector theft measures from the V-Dem datasets as a proxy.

At the outset, the average share of non-corrupt individuals was approximately 0.6, based on the average from V-Dem's 'Public Sector Corrupt Exchanges' and 'Public Sector Theft' indicators. This gradually increased to a final value of 0.8, reflecting a substantial shift in corruption levels over an average of 24 years from the start to the end of initial reformist leadership in the four-country sample. 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 $\bar{\delta}_{NC}$.

Table 1: Actual and Calibrated Parameter Values

Parameter	Initial Value	Reform Years (0–24)	Post-Reform Actual	Calibrated Final Values
θ_t^{NC} (non-corrupt share)	0.6	Evolves via transition equation	0.80	∼0.79
θ_t^{C} (corrupt share)	0.4	$1 - \theta_t^{NC}$	0.2	∼0.21
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	—	0.02
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)	—	0.07	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 section assumes that 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. The model assumes:

$$\mu_t = \begin{cases} 0.045 + (0.095 - 0.045) \cdot \dfrac{t}{24}, & \text{for } 0 \leq t \leq 24, \\ 0.02, & \text{for } t > 24\ . \end{cases} \tag{21}$$

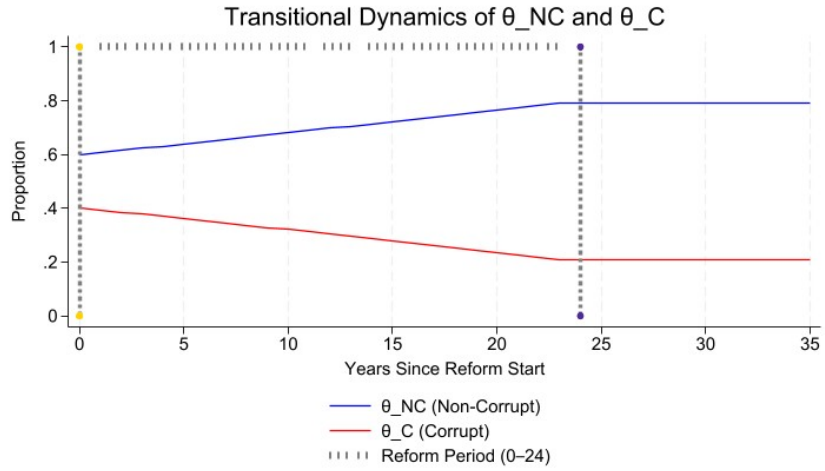
where t is the number of years since reform began. This formulation captures both habit formation and the lagged effects of educational reforms that gradually influence the composition of new entrants, gaining momentum as time increases.

Figure 3.2, Panels A and B, show the model produces the empirical path of corruption with close fit. This suggests the model not only captures the key dynamics

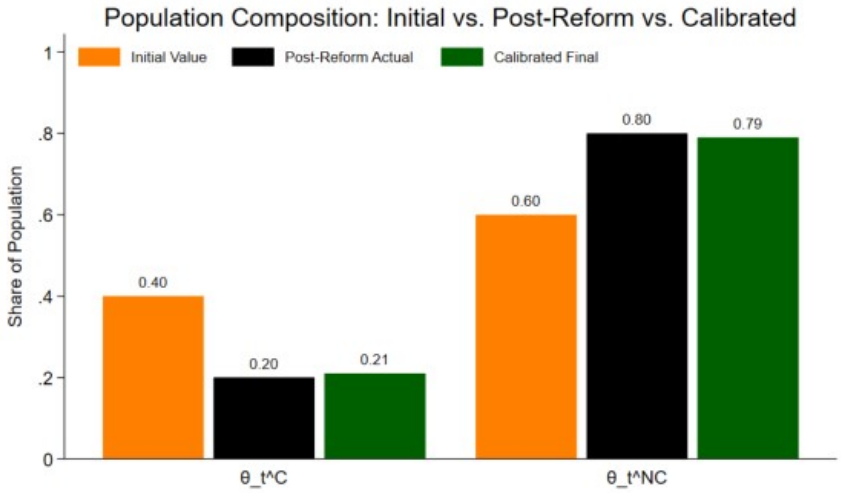
of institutional transformation but also offers a tractable framework for interpreting historical episodes and guiding future reform strategies.

How long norms take to change is a contentious issue that the calibration exercise helps illuminate. While Acemoglu and Robinson (2025) suggest norms may shift within a generation, Nunn (2023) and Guiso et al. (2016) argue it can take several. The calibration exercise suggests that the answer is possibly somewhere in the middle, sustainable change can happen within 20 to 25 years. As such, in some ways closer to Acemoglu and Robinson (2025) than Nunn (2023).

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) Model Calibration — Share of Non-Corrupt Population



(b) Model Calibration Values

Figure 3: Model Calibration Results: Panel A shows the share of non-corrupt population; Panel B shows the calibrated parameter values.

6 Tests of the model

This section empirically tests Hypotheses 1 and 2, with primary focus on Hypotheses 1. The paper finds consistent evidence supporting both.

To evaluate Hypotheses 1, the analysis uses multiple samples and methods. A key limitation is the absence of a valid external instrument, which constrains causal inter-

pretation. As Acemoglu and Robinson (2025) note, identifying causal effects in this context is inherently difficult. Accordingly, the results are interpreted as evidence of association rather than causation.

While not the main focus of the paper, this section also provides initial evidence for Hypotheses 2. The analysis examines 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).

Data

This section briefly outlines the data sources used for 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 norms. To address this, the empirical analysis relies on individual-level data from the WVS to capture ethical attitudes toward corruption among individuals and WGI indicators to capture the leadership norms.

The analysis constructs country-level indicators of ethical tolerance of corruption using WVS, 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)[16]. These variables are then averaged by country and time. These questions reflect public attitudes toward actions that undermine integrity and honest norms. From the above WVS questions, the analysis is then able to identify norms on corruption at the country level and exploit its time level variations in a panel data setting.

To test the Population–Leadership Symmetry Principle, the dataset spans 1995 to 2020, ensuring alignment of WGI and WVS coverage. The analysis draws on a broad sample of countries, 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 of 4 to 5 year gaps, I link each wave to the earliest WGI year, resulting in an average four-year gap between observations per country.

The WVS comprises five waves (1996–2020), with each country typically having

[16]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.

between 7,000 and 10,000 individual level observations per wave. 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, the paper relies on the Control of Corruption indicator from the World Bank's WGI. This metric captures the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, and the extent of state capture by elites. It closely reflects the ethical dimension of leadership emphasized in Hypothesis 1. For robustness, the paper also uses 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.[17]

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, the paper also assesses the model's predictions using an alternative source: the V-Dem dataset. This provides a complementary test, drawing on measures of particularistic spending (v2dlencmps), to captures bias in public goods provision, and of executive bribery (v2exbribe), reflecting high-level illicit exchanges within the leadership.

Test of Hypothesis 1

The paper estimates the following three equations:

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

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.

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

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

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

Here, Norm captures each country's integrity population norms, while ΔSocNorm measures changes. $\mathbf{X}$ is a vector of country and regional-level controls, described below in Table 3. δ_t represent 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, including Equations (E4) and (E5), (see below).

$\mathbf{X}$ is designed to capture variables that might correlate 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, legal origins and regional trends. These variables are discussed in the robustness section of the appendix.

Table 3: Different control sets used ($\mathbf{X}$) 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, the empirical analysis adopts a temporal separation strategy using Equations (E2), (E3) and (E4). Here, *Initial SocNorm* captures each country's lagged level of corrupt population norms. Specifically, the analysis uses lagged values of societal norms—measured 5 to 20 years prior—to examine associations with subsequent institutional quality. For robustness, several lengths of lags are used (see discussion below) as. This approach assumes that earlier values of population norms are unlikely to be shaped by current governance performance, allowing for a more credible assessment of directional influence from norms to institutions. The application of internal instruments is similar to Cingano (2014); Breunig and Majeed (2020). Additionally, the analysis controls for changes in norms (ΔSocNorm) in Equation (E3). A challenge with Equation (E3) is that it drops countries without consecutive observations. However, results remain robust to both versions.

First difference and system GMM techniques overcome reverse causality and omitted variable biases, 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). 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).

The section applies 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, assuming 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. [18]

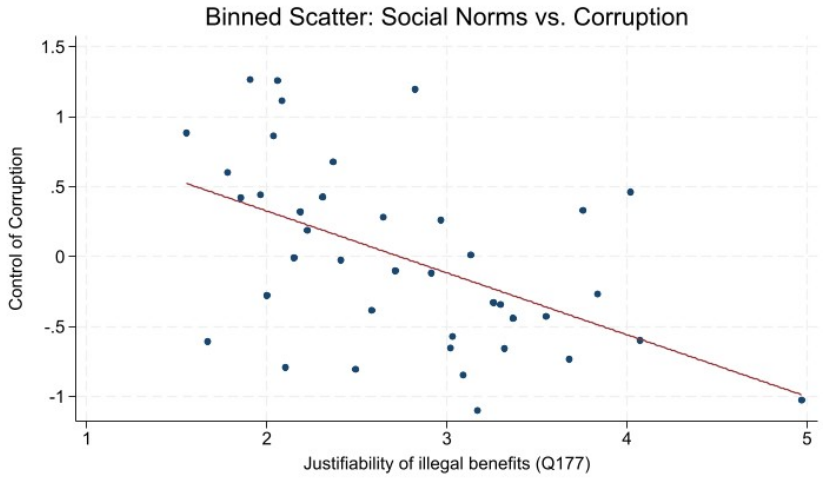
The main empirical limitation of this paper is the absence of a valid external instrument. This constrains the ability to draw causal inferences. As emphasized in Acemoglu and Robinson (2025) establishing causality in this domain is inherently challenging. Accordingly, the results should be interpreted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

Results

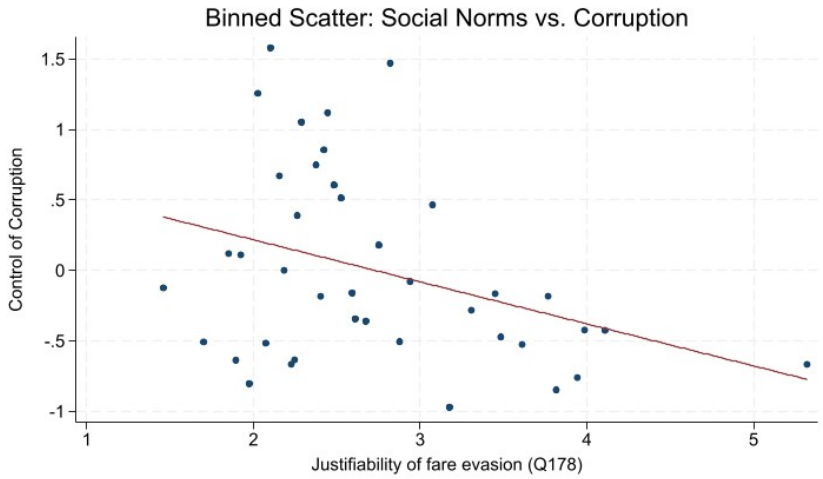
OLS

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 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.

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



(a) Justifiable to take illegal benefits



(b) Justifiable to evade fares

Figure 4: Association Between Societal Beliefs and Leadership Integrity

Tables 4 and A2 report OLS estimates of the relationship, using Control of Corruption and Rule of Law as dependent variables and proxying population characteristics with four WVS indicators of whether individuals deem dishonest behavior acceptable. Odd-numbered columns use no controls, while even-numbered columns employs control set 2. Control Set 1 yields qualitatively similar results [19], but as control set 2 additionally controls for level of development—which is a key variable—this is the preferred control set.

[19]Results available on request.

The analysis finds 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 controls, we see that economic characteristics—such as education expenditure, trade openness, and income—are also positively 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 a higher demand for integrity in leadership (Giuliano and Nunn, 2021).

Results remain qualitatively similar for all four variables on societal norms. However, as mentioned earlier, societal values on taking bribes are likely to exhibit measurement error, and as such it is dropped from further analysis. Results are also qualitatively similar for other WGI variables [20]. However, as discussed earlier, these variables are not of primary interest for here and are not considered for further analysis.

OLS estimates from Equation (E1) may suffer from reverse causality and omitted variable bias, though the control helps mitigate this possibility. 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. The section uses 2007 values as the baseline for initial societal norms, given 2007 has 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 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.

[20] Results available on request.

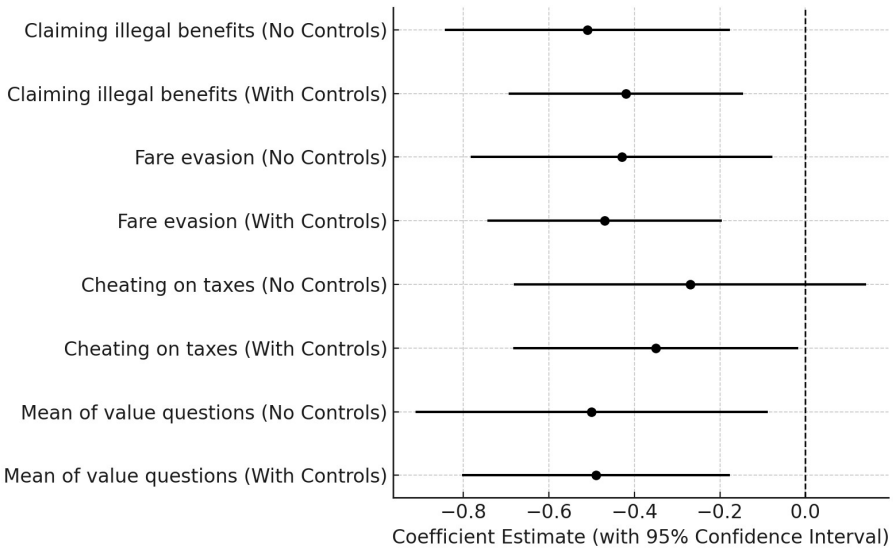


Figure 5: Effect of Initial Societal Norms on Last Leadership Characteristics

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.

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 $0.43 \times 0.87 \approx 0.37$ 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 37% of a standard deviation—a sizable effect.[21] The findings are also robust when 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 results remain robust even when addressing reverse causality with 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) have a negative and statistically significant association with leadership quality.

Again, not only are the estimated coefficients for societal tolerance of corrupt be-

[21] If I relax the requirement that countries should at least have two observations of societal values, then sample of countries increases to above 50 and results still remain robust.

havior statistically significant, they are also economically meaningful. For instance, based on the coefficient for claiming illegal government benefits in Table 6, Column (2), a one-standard-deviation increase in societal tolerance for claiming illegal government benefits (0.87) is associated with a decline of $0.40 \times 0.87 \approx 0.35$ 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, weaker societal norms are associated with a negative and statistically significant relationship with Control of Corruption across all columns. 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 represents 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, the paper finds 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 (results available on request). Since OLS and GMM yield similar estimates—and since including changes in societal norms across periods does not alter the core findings—the paper uses GMM estimates using Equation (E5) for additional robustness.

Additional Robustness

Appendix B discusses additional robustness tests in more detail. Briefly, a potential concern is that the relationship between societal norms and leadership quality may still reflect broader factors or regional forces. To address this, the paper conducts a series of robustness checks, including controls for variables that might affect both norms and leadership outcomes (e.g., electoral democracy and media freedom) and for legal origin.

This sections also constructs and controls for an index of regional institutional outcomes, excluding each country's own outcomes. The idea being that regional movements like the Arab Spring might impact social norms and leadership. Further, the section also tests Hypothesis 1 using alternative measures of leadership quality using V-Dem indicators of clientelism and executive corruption. The findings remain robust to these tests.

Results Summary

The estimates suggest that societal values around—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.

Table 4: Determinants of Corrupt Leadership (OLS)

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
Claiming illegal benefits	**-0.36*** **	**-0.29*** **						
	(0.08)	(0.06)						
Fare evasion			**-0.34*** **	**-0.34*** **				
			(0.09)	(0.07)				
Cheating on taxes					**-0.40*** **	**-0.33*** **		
					(0.11)	(0.09)		
Taking bribes							**-0.42*** **	**-0.21** **
							(0.11)	(0.10)
Education Expenditure (% GDP)		0.09**		0.13***		0.10***		0.09**
		(0.04)		(0.04)		(0.04)		(0.04)
Trade Openness (% GDP)		0.00**		0.00*		0.00*		0.00*
		(0.00)		(0.00)		(0.00)		(0.00)
Log GDP per capita		0.71***		0.74***		0.74***		0.75***
		(0.07)		(0.06)		(0.07)		(0.07)
Time = 2012	-0.33*	-0.31**	-0.43**	-0.39***	-0.46**	-0.40***	-0.37*	-0.37**
	(0.19)	(0.13)	(0.19)	(0.13)	(0.19)	(0.14)	(0.19)	(0.14)
Time = 2020	-0.37**	-0.43***	-0.45**	-0.47***	-0.54***	-0.57***	-0.45**	-0.54***
	(0.19)	(0.14)	(0.19)	(0.14)	(0.18)	(0.14)	(0.19)	(0.14)
Constant	1.40***	-6.24***	1.37***	-6.48***	1.37***	-6.52***	1.21***	-6.85***
	(0.25)	(0.68)	(0.27)	(0.64)	(0.28)	(0.70)	(0.26)	(0.75)
Observations	178	144	175	141	176	142	179	144
R-squared	0.14	0.65	0.12	0.66	0.12	0.63	0.10	0.61
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls		Yes		Yes		Yes		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*** **	**-0.42*** **						
	(0.17)	(0.14)						
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** **	**-0.49*** **
							(0.21)	(0.16)
Education Expenditure (% GDP)				0.12	0.22**	0.19*	0.18	
				(0.11)	(0.10)	(0.11)	(0.11)	
Trade Openness (% GDP)		0.00		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)
Log GDP per capita		0.54***		0.59***		0.61***		0.58***
		(0.13)		(0.12)		(0.13)		(0.13)
Constant	1.71***	-4.57***	1.52***	-5.33***	0.96*	-5.84***	1.62***	-5.07***
	(0.47)	(1.39)	(0.50)	(1.24)	(0.49)	(1.32)	(0.53)	(1.29)
Observations	44	41	43	40	44	41	44	41
R-squared	0.18	0.53	0.12	0.56	0.04	0.49	0.12	0.55
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***** (0.10)	**-0.40***** (0.08)						
Fare evasion			**-0.37***** (0.11)	**-0.42***** (0.09)				
Cheating on taxes					**-0.26**** (0.13)	**-0.34***** (0.10)		
Mean of integrity values							**-0.46***** (0.12)	**-0.47***** (0.10)
Education Expenditure		0.14** (0.06)		0.22*** (0.06)		0.19*** (0.06)		0.19*** (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*** (0.08)		0.71*** (0.08)		0.74*** (0.08)		0.70*** (0.08)
Constant	1.71*** (0.30)	-5.61*** (0.85)	1.40*** (0.33)	-6.37*** (0.78)	0.98*** (0.32)	-6.86*** (0.80)	1.57*** (0.34)	-6.10*** (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	(2) Con. Crpt	(3) Con. Crpt	(4) Con. Crpt	(5) Con. Crpt	(6) Con. Crpt	(7) Con. Crpt	(8) Con. Crpt	(9) Con. Crpt
Societal values:								
Initial: Claiming illegal benefits	**-0.56*** **	**-0.28***						
	(0.17)	(0.16)						
Change: Claiming illegal benefits		0.07						
		(0.11)						
Initial: Fare Evasion			**-0.48*** **	**-0.50** **				
			(0.16)	(0.20)				
Change: Fare Evasion				0.03				
				(0.25)				
Initial: Cheating on Taxes					**-0.36** **	**-0.43** **		
					(0.14)	(0.19)		
Change: Cheating on Taxes						0.08		
						(0.32)		
Initial: Mean of Integrity Values							**-0.58*** **	**-0.50** **
							(0.17)	(0.19)
Change: Mean of Integrity Values								0.17
								(0.20)
Education Expenditure (% GDP)		0.22**		0.41***		0.26*		0.31**
		(0.11)		(0.14)		(0.13)		(0.13)
Trade Openness (% GDP)		0.00		0.00*		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)
Log of GDP per capita		0.88***		0.57**		0.57*		0.64**
		(0.31)		(0.28)		(0.31)		(0.29)
Constant	1.85***	-8.77***	1.68***	-6.07**	1.21***	-5.79**	1.84***	-6.57**
	(0.47)	(3.16)	(0.50)	(2.50)	(0.38)	(2.78)	(0.49)	(2.71)
Control Set 2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	165	70	162	66	165	68	165	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 brevity.

Test of Hypothesis 2

Hypothesis 2: Reforms need intensity and duration to have a lasting effect: only then can they lead to a hysteresis effect. Both intensity and duration are necessary to change underlying norms.

The previous section showed that societal norms around cheating are strongly associated with leadership traits across a broad set of countries overtime, supporting Hypothesis 1. It showed how and why inertia dominates in most 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 evidence consistent with Hypothesis 2.

The analysis compares Türkiye (under Kemal Ataturk), Singapore (under Lee Kuan Yew) [22], and Pakistan — 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, corruption and education (Aziz and Mikhailova, 2016; Majeed, 2014), highlighting the importance of reform longevity for institutional durability.

To assess this empirically, the section uses an event study around key leadership transitions, providing a quasi-experimental framework to identify their impact on corruption and social norms Chalendard et al. (2023); Kwoka et al. (2016).

Direct measures of societal corruption norms are unavailable over long historical periods. For example, WVS data do not extend far enough to capture major episodes in Singapore or Türkiye, or other key episodes. However, this section presents evidence consistent with Hypothesis 2 by using proxies and tracing signs of hysteresis in corruption levels and broader social norms. The section uses broader proxies of public sector corruption and social trust, the latter viewed as an important predictor of social outcomes and 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 consistent with the hysteresis effect. It shows that in cases where reform episodes had both duration and intensity (Singapore and Türkiye), underlying norms showed evidence of a shift, with accompanying evidence of a sustained shift in public sector corruption. In the case of Pakistan, where there was reform intensity but not duration, norms did not change, and while

[22]South Korea, China, and Botswana also exemplify sustained, high-intensity leadership that produced lasting institutional reform—consistent with the hysteresis effect. By contrast, Afghanistan and Myanmar illustrate further failed reform episodes. In both cases, reforms lacked either intensity or duration. Afghanistan due to weak enforcement despite prolonged support, and Myanmar due to democratic backsliding following the 2021 military coup.

corruption improved over a short period, it was not durable. The section uses four key indicators from the Varieties of Democracy (V-Dem) dataset. To proxy corruption, the following two indicators are used: *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 treatment based on equal distribution of power (v2pepwrsoc) and views that the administration is impartial and vigorous (v2clrspct). Broadly these measures are used as proxies for social trust in the administration and fairness in the society. Higher values indicating higher social trust [23].

Event Study Regression

To test H2, the section estimates the following event study regression:

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

Y_t denotes the outcome of interest, either a corruption proxy or a social-norms proxy (as defined above). Event time is binned relative to the onset of reform leadership. ε_{it} is the error term. [24]

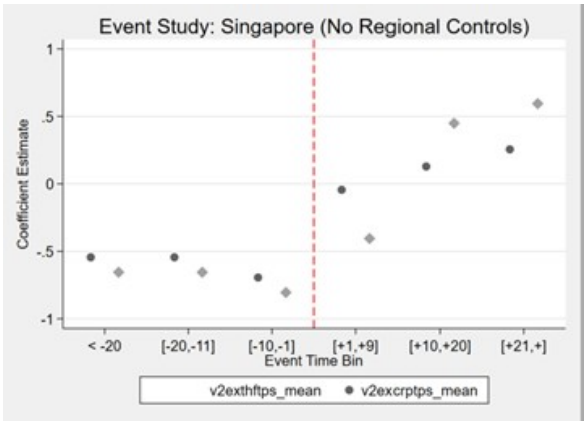
(Figure 6a–6f) presents regression estimates, with the dashed line showing the onset of a reformist regime. These results suggest that corruption and social trust improvements in Türkiye and Singapore endured well beyond the tenure of Atatürk and Lee Kuan Yew, while Pakistan's gains under Musharraf were short-lived, outcomes consistent with Hypothesis 2 [25]. In Pakistan, even though corruption indicators showed some short-run improvements, the social trust proxy remained unchanged—suggesting the regime was too short-lived to shift the underlying societal norms.

As Result 2 suggests, the changing of norms is vital for durable reforms. Given this, the analysis contrasts Singapore and Türkiye — where reform intensity and duration were long enough to change norms—with Pakistan, where they were not. This contrast provides evidence consistent with Hypothesis 2 and highlights the importance of sustained and intensive leadership in generating lasting institutional change and altering underlying norms.

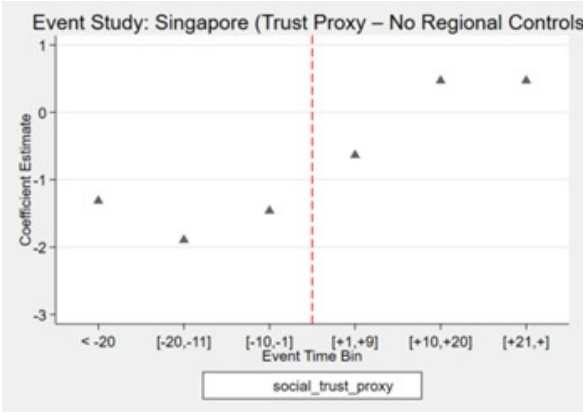
[23] Both variables are preferred over WGI variables for the longer time periods as WGI variables are not available for earlier periods.

[24] Given that Musharraf's tenure lasted roughly 8 years, event bins for Pakistan are constructed in 5-year intervals to allow clearer identification of changes, in contrast to the 10–year bins used for Singapore and Türkiye. The duration for Pakistan's study ends around 2018, 20 years after Musharraf's tenure.

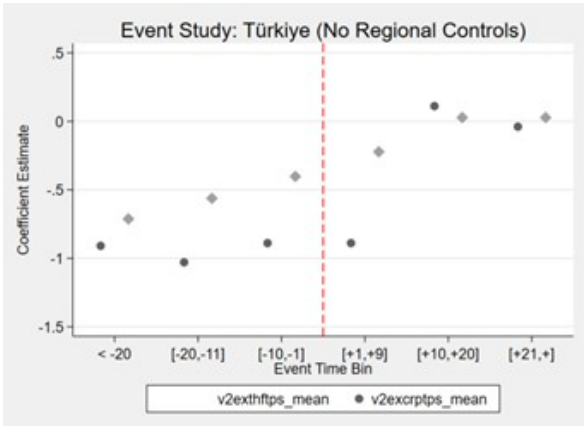
[25] Regression tables available on request.



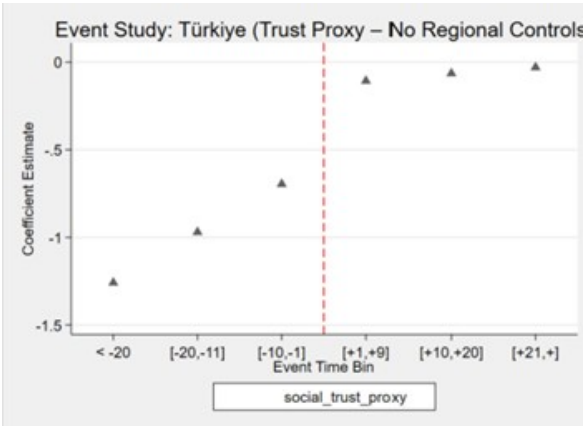
(a) Singapore. Corruption Indicators



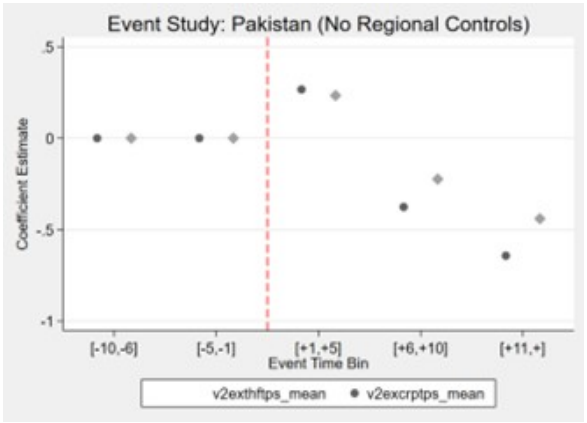
(d) Singapore. Trust Proxy



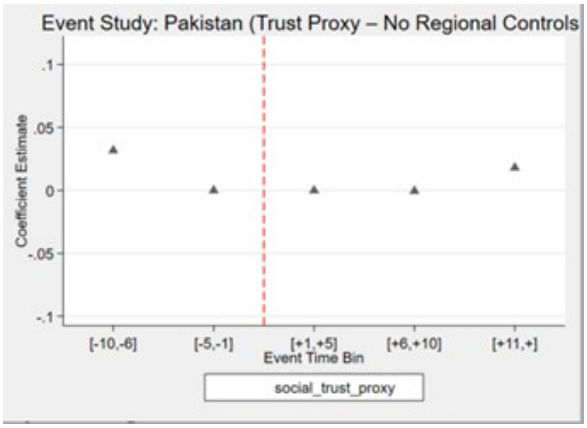
(b) Türkiye. Corruption Indicators



(e) Türkiye. Trust Proxy



(c) Pakistan. Corruption Indicators



(f) Pakistan. Trust Proxy

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

Robustness

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

Hypothesis 2 examined only a few countries. Future research could explore additional case studies to assess whether lasting reform requires norm shifts, and whether both intensity and duration matter. This could include cases of backsliding, and instances where promising initial reforms collapsed due to lost intensity or limited duration.

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, the paper proposes a novel 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 and unifying role. The paper also demonstrates that leaders pursuing regressive reforms can be especially dangerous when they combine long tenure with high reform intensity, as this increases the risk of eroding societal norms and embedding institutional backsliding.

Drawing on modelling and cross-country empirical tests, the paper documents strong alignment between societal norms and leadership integrity, consistent with the symmetry principle. It also provides evidence that durable institutional change emerges from leadership episodes that are both intense and long-lasting, consistent with the hysteresis effect. 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. It offers important policy lessons for future reformers.

The Population–Leadership Symmetry Principle and the Hysteresis Effect, though developed in the context of corruption, may also extend to other dominant societal traits. Leadership selection and norm shifts may similarly operate in religious or ideological contexts, a possibility that merits further research.

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.

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*.

Aziz, S. and Mikhailova, A. (2016). *From Banking to the Thorny World of Politics*. Quartet Books Limited, London.

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.

Gratton, G., Lee, B. E., and Yousaf, H. (2025). Bad democracy traps. *BAFFI CAREFIN Centre Research Paper*, (245).

Guiso, L., Sapienza, P., and Zingales, L. (2016). Long-term persistence. *Journal of the European Economic Association*, 14(6):1401–1436.

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.

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.

Krieger, T. (2022). Democracy and the quality of economic institutions: Theory and evidence. *Public Choice*, 192:357–376.

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.

Quah, J. S. (2022). Lee kuan yewâs role in minimising corruption in singapore. *Public Administration and Policy*, 25(2):163–175.

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 A: Table A1: Country List

Table A1: Country List by Data Sample Category

Table 5 (Shortest Sample)	**Table 5 (contd)**	**Table 4 (Full Sample)**	**Table 4 (contd)**	**Table 4 (contd)**
Andorra	Malaysia	Algeria	Hungary	Philippines
Argentina	Mexico	Andorra	India	Poland
Australia	Morocco	Argentina	Indonesia	Puerto Rico
Brazil	Netherlands	Armenia	Iran, Islamic Rep.	Qatar
Canada	New Zealand	Australia	Iraq	Romania
Chile	Poland	Azerbaijan	Italy	Russian Federation
China	Romania	Bangladesh	Japan	Rwanda
Colombia	Russian Federation	Belarus	Jordan	Serbia
Cyprus	Rwanda	Bolivia	Kazakhstan	Singapore
Egypt, Arab Rep.	Serbia	Brazil	Kenya	Slovak Republic
Ethiopia	Slovenia	Bulgaria	Korea, Rep.	Slovenia
Georgia	South Africa	Burkina Faso	Kuwait	South Africa
Germany	Spain	Canada	Kyrgyz Republic	Spain
Ghana	Sweden	Chile	Lebanon	Sweden
Guatemala	Thailand	China	Libya	Switzerland
Hong Kong SAR, China	Trinidad and Tobago	Colombia	Macao SAR, China	Tajikistan
India	Turkiye	Cyprus	Malaysia	Thailand
Indonesia	Ukraine	Czech Republic	Maldives	Trinidad and Tobago
Iran, Islamic Rep.	United Kingdom	Ecuador	Mali	Tunisia
Japan	United States	Egypt, Arab Rep.	Mexico	Turkiye
Jordan	Uruguay	Estonia	Moldova	Ukraine
Korea, Rep.	Vietnam	Ethiopia	Mongolia	United Kingdom
		Finland	Morocco	United States
		France	Myanmar	Uruguay
		Georgia	Netherlands	Uzbekistan
		Germany	New Zealand	Venezuela, RB
		Ghana	Nicaragua	Vietnam
		Greece	Nigeria	West Bank and Gaza
		Guatemala	Norway	Yemen, Rep.
		Haiti	Pakistan	Zambia
		Hong Kong SAR, China	Peru	Zimbabwe

Appendix B: Robustness Tests

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 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 [26]. 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 Public Sector Theft (v2exthftps), which measure corruption in the public sector. Results

[26]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.

are presented in Table A4, which reports system GMM estimates using covariate Set 2. Across all specifications, the core results remain robust. Additionally, the results remain robust to using 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.

Future Research

If this framework proves useful for understanding institutions, the following ideas suggest some directions for future studies.

Hypothesis 2 examined only a few countries. Future research could explore additional case studies to assess whether lasting reform requires norm shifts, and whether both intensity and duration matter. This could include cases of backsliding, as well as instances where promising initial reforms collapsed due to lost intensity or limited duration.

Similarly, while finding a global external instrument will be hard for Hypothesis 1, microdata evidence or natural experiments at the country level might be useful for future tests.

It may also be possible to expand the PA framework, either grounded in supplementary microdata evidence or by adding additional behavioral elements. There may also be additional ways to argue for behavioral symmetry.

Additional work could explore different types of population distributions, such as dispersed or bi-modal, and examine their implications for leadership.

Finally, further work is needed to understand when deviations in leadership are more likely—and under what conditions such leaders are likely to sustain longer durations in power. This could be linked to Ψ in the model.

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: Claiming illegal benefits	**-0.37***** (0.08)	**-0.29***** (0.05)				
Societal values: Fare evasion			**-0.34***** (0.09)	**-0.34***** (0.06)		
Societal values: Cheating on taxes					**-0.36***** (0.10)	**-0.29***** (0.08)
Education expenditure (% GDP)		0.06** (0.03)		0.10*** (0.03)		0.07** (0.03)
Trade openness (% GDP)		0.00*** (0.00)		0.00** (0.00)		0.00* (0.00)
Log of GDP per capita		0.69*** (0.05)		0.72*** (0.05)		0.74*** (0.06)
Constant	1.36*** (0.24)	-6.01*** (0.57)	1.30*** (0.26)	-6.27*** (0.54)	1.23*** (0.27)	-6.48*** (0.60)
Observations	178	144	175	141	176	142
R-squared	0.14	0.71	0.11	0.72	0.10	0.67
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) Con. Crpt	(3) Con. Crpt	(4) Con. Crpt	(5) Con. Crpt	(6) Con. Crpt	(7) Con. Crpt
Societal values: Claiming illegal benefits	**-0.56*** **	**-0.33** **				
	(0.17)	(0.13)				
Societal values: Claiming illegal benefits = D,		0.06				
		(0.08)				
Societal values: Fare evasion			**-0.48*** **	**-0.50*** **		
			(0.16)	(0.14)		
Societal values: Fare evasion = D,				0.10		
				(0.20)		
Societal values: Cheating on taxes					**-0.36** **	**-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	1.85***	-4.70	1.68***	-4.02	1.21***	-3.68
	(0.47)	(2.84)	(0.50)	(2.84)	(0.38)	(3.86)
Observations	165	67	162	63	165	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	(2) PS Theft	(3) PS Theft	(4) PS Theft	(5) PS Theft	(6) PS Theft	(7) PS Theft	(8) PS Theft	(9) PS Theft
Societal values:								
Claiming illegal benefits	**-0.46*** **	**-0.24***						
	(0.13)	(0.13)						
Claiming illegal benefits = D,		0.01						
		(0.06)						
Fare Evasion			**-0.39*** **	**-0.35** **				
			(0.13)	(0.14)				
Fare Evasion = D,				-0.06				
				(0.18)				
Cheating on taxes					**-0.32** **	**-0.33** **		
					(0.14)	(0.14)		
Cheating on taxes = D,						0.06		
						(0.26)		
Mean of different Qs							**-0.49*** **	**-0.38*** **
							(0.14)	(0.14)
(mean) = D,								0.08
								(0.14)
Education Expenditure (% GDP)		0.01		0.16		0.08		0.09
		(0.14)		(0.14)		(0.13)		(0.12)
Trade Openness (% GDP)		0.00		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)		(0.00)
Log of GDP per capita		0.78***		0.60***		0.60***		0.66***
		(0.21)		(0.16)		(0.18)		(0.16)
fn_Q180								
Constant	3.96***	-4.56**	3.79***	-3.11*	3.49***	-2.93*	3.99***	-3.50**
	(0.37)	(2.16)	(0.40)	(1.55)	(0.36)	(1.71)	(0.41)	(1.51)
Observations	162	67	159	63	162	65	162	67
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of country_id	54	37	53	36	54	36	54	37

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