THE MYSTERY OF COOPERATION

A Look into Costly Signaling and Cross-Cultural Coordination

Gimenez, Gabriela E.
ECON 420 – Economics Colloquium

Abstract

This paper develops a theory of cross-cultural cooperation rooted in a simple insight. Cooperation depends on players agreeing on the rules of the game. As on a playground, children of different maturity levels can play low-contact games with limited harm, while high-contact games produce unequal injury over time even without intent. The same logic applies to cooperation across cultural and institutional distance. Trade becomes fragile when agreed-upon rules impose unequal risks. The paper argues that diasporic individuals seek cooperation across cultural distance by investing in costly, relationship-specific signals that reduce uncertainty and make exchange possible. When host institutions cannot reliably interpret or value these signals, the costs of cooperation fall asymmetrically on diasporic actors, creating persistent hazards and governance misalignment. Over time, this asymmetry erodes cooperative equilibria even when gains from trade remain positive. The theory explains why cooperation under asymmetric signaling often begins, persists unevenly, and collapses intertemporally rather than failing at entry.

Acknowledgements

I am grateful to my cohorts, to Dr. Herbener, Dr. Fuller & Dr. Truitt for their guidance, and to my friends for patiently listening to my endless thoughts on transaction costs.

I. Welcome to the Playground

Games are among our earliest experiences with mutually beneficial exchange. Even children who dislike each other will still play the same game as long as they agree upon fair rules to play by. Rules are the formally agreed-upon expectations about fairness, conduct, and outcomes. The rules of the game arise spontaneously in the environment in which they develop, coordinating behavior without central design (Hayek 1960). Rules work when they align incentives, which makes the prize of winning worth more than the cost of playing. This is why the specific features of players of similar maturity do not matter. Because if the game is fair and the rules hold, the prize outweighs both discrimination and disagreement.

Are all players in a game equal? No. Each player has a distinct comparative advantage, unique information, and opportunity costs for engaging in a game. However, fair games can happen with parties with different features; it is when you get parties at different maturities to play that the game can become risky. A five-year-old playing chess with a twelve-year-old may suffer only emotional injury, but the overall benefit of the game remains positive. Yet, the same pairing in a wrestling match would be disastrous. In low-contact games, this asymmetry produces little long-term harm. In high-contact games, it can become dangerous. Why is this true? Because in this kind of scenario, the younger child compensates by expending effort beyond their capacity. The older child may not even notice the difference between strengths or may deliberately exploit it. Either way, the dynamic creates a structural mismatch. Whether rough play is intentional or not, it disproportionately harms the smaller child.

This same phenomenon can also occur between differing cultures and their institutional norms. Cultures differ in institutional maturity, norms, expectations, and signaling repertoires. Cross-cultural exchange can either feel like differing-aged children playing chess or like

wrestling across weight classes. This is why wrestling prohibits matches outside a weight class: the expected harm is structurally unequal. As the institutional gaps widen, the cost of cooperation rises, and the equilibrium becomes fragile. This leads to the central questions of this paper:

How do humans sustain cooperation across cultural differences when one group is more institutionally mature than the others? Why do some exchanges thrive or collapse over intertemporal exchange relationships?

Cooperation is a rule-governed equilibrium that emerges when expected gains from trade exceed the costs of participation and enforcement. In cross-cultural settings, cooperation depends not only on shared rules but on the ability to interpret them. When interpretive burdens fall asymmetrically, diasporic individuals bear disproportionate costs, making cooperative equilibria fragile in high-stakes interactions.

This paper argues that cross-cultural cooperation is sustained through costly signaling that reduces social and institutional distance. When such signaling constitutes a specific investment but cannot be reliably observed, interpreted, or reciprocated by host institutions, the costs of cooperation fall asymmetrically on diasporic actors. Over time, this asymmetry erodes cooperative equilibria even when mutual gains from trade exist.

The paper first answers what cooperation is by drawing on the literature to establish the universal foundations for cooperation. It then asks: What is culture? We will use North, Sowell, de Soto, and Leeson as the foundation for establishing culture. Next, it identifies the literature gap: cross-cultural cooperation is asymmetrically costly, yet this mechanism is not modeled. It then develops the mechanism: diasporic signaling as a specific investment under asymmetric observability (Leeson + Williamson + KCA). The framework is applied to Black Caribbean and

African diasporic communities in U.S. professional institutions, drawing on audit studies, name experiments, and ethnographic work. Finally, it draws implications: who wins, who loses, and why cooperation collapses *intertemporally* rather than in a single interaction, and closes by arguing that this is not a zero-sum game.

This paper is a theory paper that develops a mechanism explaining why cross-cultural cooperation can erode over time even when gains from trade exist, by modeling diasporic signaling as a costly, specific investment under asymmetric observability in repeated interaction.

II. What is Cooperation?

Mutual gains exist. Mises emphasizes social cooperation as the outcome of the division of labor: Individuals specialize because doing so increases both productivity and mutual gains (Mises 1998). In Human Action, chapter 8, Mises describes Ricardo's law of comparative advantage, which shows that even when one party is absolutely more productive in every task, it still remains mutually beneficial to exchange (Ricardo 1951). This is because each party must economize on production costs, therefore leading them to produce only the thing most valuable, because the opportunity cost of the second most highly valued thing is too high in comparison. Cooperation can be beneficial even when groups are unequal in capabilities. What this paper emphasizes is how rules structure the sharing of gains. The question becomes whether rules prevent the inequality from turning into "high-contact" cooperation into extraction.

Hayek argues that rules and institutions emerge as spontaneous orders that economize on dispersed knowledge (Hayek 1960). Rules can emerge and evolve without a central authority to design or enforce them. This is echoed in Smith's *Theory of Moral Sentiments,* where he provides a framework demonstrating how expectations of propriety and fairness become socially

stabilized (Smith 1759). Buchanan and Tullock frame politics and collective choice as the choice of "rules of the game," under which individuals pursue their own interests (Buchanan & Tullock 1962). Playground rules let strangers cooperate. Constitutional and moral rules do the same at scale.

Cooperation in economies can be understood as a rule-governed equilibrium rather than sentiment: like games, exchange requires shared expectations about permissible conduct and the distribution of gains. These rules emerge as locally efficient solutions to coordination problems and persist when they align incentives, so participation yields a net advantage relative to exit or conflict.

Coase's "The Nature of the Firm" explains that firms arise when the cost of using the price mechanism (search, bargaining, enforcement) is high (Coase 1937). "The Problem of Social Cost" shows that real-world exchange is shaped by transaction costs rather than just legal entitlements (Coase 1960). Even when entitlements are clarified, bargaining and post-judgment adjustment may not proceed smoothly because of strategic behavior and frictions that arise in real dispute resolution (Farnsworth 1999). This is the gap between Coasean possibility and lived institutional reality that motivates governance-and-hazards analysis. Some institutional knowledge is tacit rather than codified; people follow rules they cannot fully state (Polanyi 1966). When "what counts" is tacit, interpretability becomes the bottleneck—setting up why hazards persist even when everyone is trying.

Williamson develops an entire system of analysis to answer a specific question: what could block the gains from exchange? (Williamson 1985, 2010). Whenever people cooperate, whether in a simple spot exchange or within a firm, they take actions to capture some of those gains from trade. Because people both take these actions and others can anticipate they'll take these actions,

sometimes the gains from trade do not arise in the first place. For example, a party could say, "I will not engage in cooperation with you because I anticipate that you are going to do something that makes me worse off later." That anticipation blocks the gains from trade, so Williamson asks what sorts of actions people take to overcome that hurdle, connecting to the "hospitality tradition" idea: practices that look inefficient or wasteful often persist because they solve a real problem. Many behaviors we treat suspiciously are undertaken because they enable cooperation under hazard: without them, parties would not engage in exchange at all. Williamson shows how different markets develop by examining the range of payment and contractual forms as attempts to overcome this hurdle—using asset specificity, opportunism, and governance alignment as key concepts (Williamson 1985, 2010). Specific investments can generate appropriable quasi-rents and hold-up risk. Leading to governance responses (including integration) that are shaped by that hazard (Klein, Crawford, and Alchian 1978). This provides the backbone for treating diasporic signaling as a form of specific investment that can be expropriated or undervalued.

III. What is Culture?

Douglas North takes transaction cost reasoning to the historical record (North 1991). His question: why did trade in small tribal communities look one way, and then as cities arose, trade and production organization looked entirely different? The animating impulse parallels Williamson: different environments pose different challenges to capturing the gains from trade. As cities emerge and markets expand, people develop new ways of organizing production and exchange to capture those gains under new constraints. The historical analysis becomes a progression of institutional changes that arise to maximize cooperation as transaction costs and opportunities evolve. Sowell, in his work "Ethnic America," examines how ethnic groups differ in cultural patterns—values, time horizons, social norms—that shape economic performance

(Sowell 1981). Culture here is a set of accumulated responses to historical constraints and opportunities.

Culture is institutional memory. Culture is the accumulated set of beliefs, practices, memories, and values that communities choose to preserve and transmit. Economically, culture functions as a form of inherited capital as a repertoire of rules, heuristics, stories, and expectations that reduce uncertainty and structure cooperation. Cultures protect the most efficient means of value creation given a particular environment. De Soto shows how mature institutions embody cultural memory through formal property rights, registries, and legal codes (de Soto 2000). For insiders, these codified forms feel "natural" since they no longer remember the process by which these rules were generated. For outsiders, the absence of that institutional memory makes the system opaque: they see the rules but not the history that makes them legible. Leeson's work on medieval trade and social distance shows that when groups share norms and interpretive frames, they can sustain self-enforcing exchange. When social distance is high, self-enforcement becomes costly (Leeson 2006, 2008). This provides the bridge between culture as institutional memory and the mechanism this paper will utilize: norm distance increases enforcement cost and raises the hazard that signals will be misread. Cross-cultural cooperation differs because *signal-translation costs* vary across cultural-institutional environments. This paper contributes a theory of cross-cultural cooperation as an intertemporal equilibrium under asymmetric signal observability.

IV. The Problem of Cross-Cultural Cooperation

A baseline explanation for unequal treatment in cooperative settings is taste-based discrimination. Becker argues that some decision-makers derive disutility from interacting with particular groups, leading them to behave as if the "price" of hiring or cooperating were higher,

even when productivity remains unchanged (Becker 1957). Cooperation can fail even when gains from trade exist because some players perceive the prize as smaller than it actually is. Statistical discrimination offers a related account. Under uncertainty, decision-makers may rely on observable group membership as a proxy for unobserved productivity or trustworthiness (Phelps 1972). This framework helps explain why signaling becomes necessary in the first place: when hosts rely on coarse group-level proxies, diasporic actors face incentives to overinvest in counter-signals. Together, Becker and Phelps clarify why unequal outcomes persist, but they do not model diasporic signaling as a specific investment subject to asymmetric observability and intertemporal depletion.

Empirical work reinforces this limitation. Audit and résumé studies consistently show that identical credentials do not yield identical outcomes. Applicants with the same qualifications receive different callback rates depending on the racial associations of their names. This indicates that signals do not clear the market symmetrically (Bertrand and Mullainathan 2004). The same move in the game is judged differently depending on who makes it. Field experiments in low-wage labor markets further show that penalties persist even when observable qualifications are held constant (Pager, Western, and Bonikowski 2009). These studies establish the fact pattern of misrecognition and discounting. However, they do not provide a transaction-cost or governance account of why repeated interaction fails to stabilize cooperation rather than correcting these disparities over time.

Related literature points to problems of interpretability. Employers and institutions routinely evaluate demeanor, speech, and other "soft skills" as indicators of productivity or trustworthiness, even though these signals are culturally coded and locally interpreted (Moss and Tilly 1996). This creates a setting in which sincere effort does not guarantee correct

interpretation. Polanyi's concept of tacit knowledge sharpens this point. Much of what institutions enforce cannot be fully articulated; actors follow rules they cannot entirely state (Polanyi 1966). As a result, outsiders may invest heavily in translating themselves into acceptable forms without any assurance that the translation will be recognized. The problem is not only that signals are costly, but that the knowledge required to interpret them is partly tacit, making observability and legibility binding constraints on cooperation.

Work on diaspora and institutional navigation provides further context for this dynamic. West Indian immigrants in the United States engage in strategic identity presentation and boundary management in response to how host institutions categorize Blackness (Waters 2001). Caribbean migration histories and institutional contexts shape how these strategies develop and persist (Foner 2001). Institutional interfaces, including welfare state programs and racialized policy categories, further condition how groups are treated and how they respond (Lamb 1999). Evidence on differences in immigrant incarceration is often interpreted through selection and behavioral channels, which is relevant here not because incarceration is the outcome of interest, but because it illustrates how diasporic groups rationally invest in reputation and rule-following behavior when they anticipate institutional filtering (Butcher and Piehl 2007).

Access to cooperation is also shaped by networks. Hiring and new opportunities often occur through social ties, which shape which signals are noticed and how they are judged (Royster 2003). This strengthens the central point: diaspora members may pay to send signals, but it is the host's system that decides if those signals get seen at all.

Institutional economics approaches the problem from yet another angle. North emphasizes that institutions evolve as accumulated solutions to recurring coordination problems, embedding rules and expectations that insiders come to treat as natural (North 1991). De Soto shows how

these inherited practices are eventually formalized through property rights and legal systems, converting local knowledge into enforceable structures (de Soto 2000). Sowell similarly treats culture as a durable stock of responses shaped by historical constraints rather than as a set of arbitrary preferences (Sowell 1981). These accounts support the conception of culture as institutional memory, but they do not specify how that memory alters the cost of cooperation for outsiders.

The literature on self-enforcing exchange addresses cross-group cooperation more directly. Leeson shows that when groups do not share norms or interpretive frameworks, sustaining exchange requires costly investments that reduce social distance and make commitment credible (Leeson 2006; Leeson 2008). This work explains why outsiders invest heavily in norm-bridging behavior, but it does not fully account for what happens when those investments fail to translate into durable cooperation.

Transaction cost economics supplies another essential piece. Williamson centers his analysis on the hazards that arise when parties make relationship-specific investments under conditions of opportunism (Williamson 1985; Williamson 2010). When governance structures fail to align incentives, the gains from trade may never materialize. Klein, Crawford, and Alchian formalize this logic by showing how asset specificity generates appropriable quasi-rents that invite hold-up and value capture (Klein, Crawford, and Alchian 1978). These theories explain why cooperation can break down under specific investment, yet they have not been applied to cultural signaling as a systematic source of asymmetry.

What remains missing is a framework that brings these strands together. Existing literatures document discrimination, culturally coded evaluation, institutional inheritance, norm distance, and governance hazards in isolation. They do not model cross-cultural cooperation as an

intertemporal equilibrium in which diasporic signaling functions as a costly, specific investment that is unevenly observed, imperfectly interpreted, and vulnerable to appropriation.

If diasporic signaling operates as a specific investment under asymmetric observability, several patterns should follow. First, diasporic individuals should systematically overinvest in signals treated as culturally neutral—such as credentialing, presentation, and code-switching—relative to the host-group baseline. This overinvestment reflects the need to clear a higher evidentiary bar, not the intrinsic productivity of the signals themselves. Empirical work on soft skills, résumé screening, and name-based evaluation provides anchors for this expectation, even though it does not model the mechanism directly (Moss and Tilly 1996; Bertrand and Mullainathan 2004; Pager, Western, and Bonikowski 2009).

Second, returns to identical signaling effort should vary persistently across institutions. Because evaluation relies on tacit knowledge and inherited institutional memory, the same investment does not convert into trust or opportunity at uniform rates. What counts as professionalism or fit is partly unstated and locally learned, so identical signals may be recognized in one setting and discounted in another (Polanyi 1966; North 1991).

Finally, the breakdown of cooperation should be intertemporal rather than immediate. Cooperation does not fail in a single interaction. It erodes as diasporic actors are required to repeatedly reinvest in costly, non-transferable signals whose returns are uncertain and are often appropriated. Over time, the expected benefits of cooperation fall below the cost of continued signaling, making withdrawal or parallel institutional development rational. This erosion follows directly from the logic of asset specificity, hold-up, and governance misalignment, rather than from sudden conflict or overt refusal to cooperate (Williamson 1985; Klein, Crawford, and Alchian 1978; Leeson 2008).

Cross-cultural cooperation is therefore asymmetrically costly. Diasporic communities must continually invest in signals that are costly, specific, and often misread, while existing literatures treat discrimination, culture, and transaction costs in isolation. This paper proposes a unified mechanism that treats diasporic signaling as a specific investment made under conditions of asymmetric observability. By synthesizing Williamson's analysis of hazards and governance, Leeson's account of norm distance and costly signaling, and the Klein–Crawford–Alchian logic of appropriable rents, the framework explains why cooperation can emerge, persist for a time, and yet erode intertemporally when the costs of signaling and enforcement fall disproportionately on one side. The next section makes this mechanism explicit by specifying the rules of the game that sustain cooperation or, eventually, break it down.

V. What are the Rules of the Game?

Cross-cultural cooperation can be treated as a repeated game with shared surplus but uneven interpretability. The gains from exchange exist, but they are only realizable when participants can understand the rules that govern play and when the costs of enforcing those rules are not concentrated on one side. This section identifies the basic conditions under which cooperation is privately rational and stable, and those under which it becomes fragile or predatory.

First, potential mutual gain must exist; without surplus, there is no reason to cooperate in the first place. Second, the relevant rules—formal or informal—must be intelligible to participants. This does not require perfect agreement, but it does require a shared capacity to recognize what counts as compliance, defection, and repair. Third, baseline capacity parity must be sufficient for meaningful participation. When parties are mismatched in institutional maturity or bargaining position, participation remains possible, but the interaction becomes "high-contact," in the sense that small errors, misunderstandings, or opportunism can cause asymmetric harm.

Rules matter because they align incentives. When the expected private payoff from following the rules exceeds the payoff from defection, cooperation becomes rational rather than sentimental. In settings characterized by distance and uncertainty, this alignment is rarely automatic; it is produced through safeguards that lower transaction costs, reduce interpretive noise, and raise the cost of opportunism (Williamson 1985; Williamson 2010; Leeson 2008). In other words, rules do not merely constrain behavior. They make exchange feasible.

A cooperative equilibrium is stable when two conditions hold. The first is perceived fairness: participants must regard the ruleset as tolerably fitted to their capabilities and risks. The second is enforceability: defection must be detectable and sufficiently costly, whether through formal sanctions or informal mechanisms that make opportunism unprofitable (Leeson 2008; Williamson 1985). When these conditions hold, repeated dealings generate trust as an equilibrium outcome rather than a moral achievement.

Cooperation becomes fragile when high-contact asymmetry is paired with weak safeguards. Where one side bears most of the cost of making itself legible—through translation, credentialing, code-switching, or other costly signals—while the other side retains discretion over recognition and reward, the interaction contains a built-in extraction margin. In transaction-cost terms, costly, relationship-specific investments create hazards when they can be discounted or appropriated without reciprocal obligation (Klein, Crawford, and Alchian 1978; Williamson 1985). Under those conditions, cooperation does not typically collapse in a single round. It erodes as repeated reinvestment becomes unprofitable.

When interpretability is limited and enforcement is host-controlled, actors may move beyond signaling and into bonding. In transaction-cost terms, “hostages” are commitments that raise the cost of defection by making future losses contingent on present conduct. In cross-cultural

settings, these hostages are often asymmetric: the outsider concentrates downside risk to make cooperation possible under uncertainty. Such bonding can temporarily stabilize exchange, but it can also deepen exposure by increasing specificity and dependence, accelerating the intertemporal erosion when reciprocity fails (Williamson 1985; Klein, Crawford, and Alchian 1978).

VI. What Game are we Playing?

Up to this point, the paper has shown why cooperation can fail even when gains from trade exist. This section does something more precise. It defines the strategic environment—the "game"—in which diasporic cooperation with host institutions takes place. Doing so clarifies the kind of failure under examination. It also explains why the same mechanism can generate temporary success followed by long-run erosion. The breakdown described here is not the result of ignorance, animus, or a one-shot miscalculation. It arises from the structure of repeated interaction under asymmetric interpretability.

The interaction at issue is not a one-time exchange. Diasporic individuals and host institutions engage repeatedly through job searches, evaluations, promotions, client relationships, and reputation building. Each interaction shapes expectations in later rounds. This places the analysis squarely within a repeat-play environment, where trust, reputation, and the ability to interpret behavior matter for sustaining cooperation over time (Williamson 1985; Leeson 2008).

There are three analytically distinct players in this game. The first is diasporic individuals, who seek access to cooperation in the form of employment, advancement, and institutional credibility, and who bear the cost of translating norms and signals across cultural distance. The second is host institutions, which define the dominant rules of professionalism, evaluation, and legitimacy, and are often unaware of the full costs borne by outsiders in attempting to comply

with them. The third is the institutional environment itself, composed of informal norms, evaluative heuristics, and tacit expectations that structure how behavior is interpreted. This environment is not neutral; it shapes observability and payoffs, thereby affecting strategic behavior (Polanyi 1966; North 1991).

The object of the game is sustained cooperation rather than mere entry. Entry into an institution is only the first hurdle. Long-run cooperation requires being trusted, understood, and evaluated as legitimate over time. This distinction is central to the paper's argument because initial success does not imply stability. The theory developed here is designed to explain precisely why cooperation can begin and yet still unravel.

In this framework, cultural distance is not treated as a difference in preferences. It is an informational problem. Host institutions have incomplete information about outsiders' intentions, competence, and trustworthiness. Diasporic individuals, in turn, have incomplete information about which signals will be recognized, valued, or ignored. This informational wedge mirrors the conditions that make governance and safeguards necessary in transaction cost economics, but here the wedge is driven by norm distance rather than by technology or asset characteristics (Williamson 1985; Leeson 2008).

Because cooperation is valuable but risky, diasporic individuals invest in signals intended to reduce uncertainty and make exchange self-enforcing. These signals include dress, speech, credentialing, demeanor, name choices, and code-switching. Their purpose is not expressive but instrumental. They are used to demonstrate commitment, reduce perceived risk, and stabilize repeated interaction under distance (Leeson 2008). In this environment, such behavior is rational.

The asymmetry lies not in the existence of signaling but in who bears its cost. Diasporic actors pay in time, money, emotional labor, and identity management. Host institutions typically

do not incur the cost of interpreting those signals correctly. The result is an asymmetric transaction-cost burden rather than a symmetric bargaining problem.

Despite this asymmetry, cooperation often begins. This point is essential. The theory does not predict immediate failure. Early rounds may succeed because some signals are recognized and some evaluators are attentive. Formal credentials can also partially substitute for trust. Initial gains from trade exist on both sides. Labor is supplied, output is produced, and relationships form. This is why the puzzle exists at all.

Early cooperation, however, can be misleading. For cooperation to stabilize, signals must be observable, legible, and mutually interpretable. When host institutions cannot fully articulate their own evaluative criteria, even sincere compliance may not be credited. Polanyi's insight into tacit knowledge is crucial here. Much of what institutions enforce cannot be fully stated, which means that cooperation can occur without resolving the underlying informational asymmetry (Polanyi 1966).

Because signals are often undervalued, misread, or ignored, diasporic individuals must reinvest repeatedly to maintain cooperation. Signaling is not a one-time act. Code-switching, demeanor management, and credential accumulation require continuous effort. This is where the interaction becomes high-contact rather than low-contact.

These signals are costly, non-transferable, and more valuable inside the host institution than elsewhere. They therefore constitute specific investments in the sense used by transaction cost economics (Klein, Crawford, and Alchian 1978; Williamson 1985). When returns to these investments are uncertain or appropriated, the marginal cost of continued cooperation rises while expected benefits fall. The cooperative equilibrium becomes fragile.

Importantly, collapse is intertemporal rather than immediate. Cooperation does not fail in a single interaction. It erodes as repeated, unreciprocated investment drains resources. Trust is not reinforced at the same rate it is demanded. Signaling becomes privately irrational. This is the economic breaking point implied by the logic of asset specificity and governance misalignment.

This framework clarifies what the paper does not claim. It does not reduce outcomes to taste-based discrimination, statistical inference errors, or prejudice alone (Becker 1957; Phelps 1972). The literatures identify patterns of unequal treatment. They do not explain how cooperative relationships unravel over time when gains from trade exist, and actors behave rationally.

In the game described here, even good-faith hosts, rational actors, and successful early cooperation can produce long-run failure. This occurs when signal translation is asymmetric, and governance is misaligned.

VI. Chess or Checkers?

The purpose is not to establish that discrimination exists, nor to catalog inequities. Rather, it is to show how cooperation erodes over time, even when rational actors on both sides are attempting to cooperate, using well-established empirical and ethnographic work as anchors for each step of the mechanism.

The application focuses on Black Caribbean and African diasporic communities navigating professional institutions in the United States. This setting is analytically powerful because it combines repeated interaction, high stakes, substantial cultural distance, and well-documented signaling behavior. These features make the underlying mechanism visible rather than latent.

The case is chosen for theoretical clarity rather than moral salience. It exhibits high signaling intensity, as diasporic actors invest heavily in dress, speech, credentialing, demeanor, naming practices, and other markers of professionalism. The cost of this signaling is borne primarily by

the diasporic individual rather than by the host institution. At the same time, host-side interpretation is often mediated through culturally coded heuristics, so that signals are filtered in ways that do not reliably reflect their cost or intent. Finally, the long-run outcomes in this setting display instability rather than convergence, taking the form of attrition, burnout, wage gaps, and stalled advancement. Few empirical contexts make the micro-mechanism observable.

Diasporic individuals invest in signals rated as culturally neutral—grooming, dress, résumé formatting, credentialing, speech modulation, tone, and code-switching. These ongoing adjustments require time, money, and self-monitoring.

Empirical work documents these investments clearly. Waters shows that West Indian immigrants adopt respectability strategies explicitly designed to counter prevailing stereotypes in labor markets, reshaping presentation and identity in response to host expectations (Waters 2001). Butcher and Piehl show that Caribbean immigrants engage in selection and self-presentation strategies that reduce institutional penalties, consistent with anticipatory signaling under perceived hazard (Butcher and Piehl 2007).

From the perspective of transaction cost economics, these behaviors fit the profile of specific investment. They are costly, non-transferable, and context-dependent, yielding higher value within the host institutional environment than elsewhere. As such, they are vulnerable to undervaluation or appropriation in the sense described by Klein, Crawford, and Alchian and by Williamson (Klein, Crawford, and Alchian 1978; Williamson 1985).

Despite these significant investments, host institutions frequently misinterpret or discount diasporic individuals' signals. The core mechanism driving erosion of cooperation is that evaluation relies on culturally coded heuristics, informal standards of fit, and tacit notions of professionalism, causing sincere and costly efforts to go unrecognized or undervalued.

Empirical evidence supports this claim. Moss and Tilly show that employers routinely interpret soft skills as productivity signals despite their deep cultural coding, treating them as universal rather than locally grounded indicators (Moss and Tilly 1996). Bertrand and Mullainathan demonstrate that identical résumés receive different callback rates based solely on racially associated names, indicating that signals do not clear the market symmetrically (Bertrand and Mullainathan 2004). Pager, Western, and Bonikowski show that callback penalties persist even when observable credentials are held constant, reinforcing the idea that signal interpretation, not just signal provision, is the binding constraint (Pager, Western, and Bonikowski 2009).

In Leeson's framework, signals can sustain cooperation only when they are observable, legible, and mutually interpretable. This case shows that legibility fails systematically, even when observability is high, and effort is sincere (Leeson 2008).

Increased effort does not resolve the asymmetry. Even when diasporic actors fully comply with expectations, signals may be misinterpreted, effort may not be credited, and trust may still lag. The informational gap persists despite compliance.

Pager's work shows that even rehabilitated Black men continue to face substantial penalties, suggesting that effort alone does not reset institutional expectations (Pager 2007). Royster demonstrates that hiring networks interpret identical signals differently depending on group membership, shaping access before formal evaluation occurs (Royster 2003). Fryer and Levitt show that naming practices function as endogenous signals shaped by institutional response, reinforcing feedback loops rather than clearing them (Fryer and Levitt 2004). Together, these findings point to persistent asymmetric information rather than a temporary mismatch.

Because signals must be sent continually, costs build while benefits remain uncertain. Marginal returns diminish over time. Eventually, continued cooperation becomes privately irrational, though it remains socially beneficial.

Empirical work captures this slow erosion. Waters documents second-generation disillusionment and identity renegotiation as initial optimism gives way to sustained institutional friction (Waters 2001). Anderson shows how alternative norms and parallel systems emerge when trust in mainstream institutions erodes, not through sudden rejection but through gradual disengagement (Anderson 1999). The pattern is not collapse but thinning: cooperation weakens as reinvestment becomes unprofitable.

The metaphor introduced earlier now does analytical work. In low-contact settings analogous to chess, asymmetry exists, but harm is limited, and cooperation remains positive-sum. In high-contact settings, such as wrestling across weight classes, asymmetry causes real harm, and rules designed for parity fail in mismatches. U.S. professional institutions often assume chess-like interaction, while diasporic actors experience something closer to wrestling. The difference lies not in intent but in structure.

Avner Greif's study of the Maghribi traders provides a useful contrast. In that setting, norms were shared, signals were mutually interpretable, and enforcement was endogenous, allowing cooperation to stabilize despite distance (Greif 1993). Greif's case illustrates the mechanism's prediction when interpretability is high. The present application shows what happens when it is not.

VII. Who Wins, Who Loses, Who Tells the Story?

Within this framework, host institutions benefit from asymmetric signaling in ways that are largely invisible at the point of interaction. First, institutions receive risk reduction without bearing the corresponding costs. Diasporic actors engage in extensive self-monitoring, self-discipline, and overinvestment in compliance with institutional norms. This behavior lowers perceived risk for employers and professional gatekeepers, yet it does not require reciprocal investment in interpretation, mentorship, or protection. The institution benefits from increased predictability while externalizing the cost of producing it.

Second, host institutions retain option value. Because the cost of signaling is borne by the diasporic actor, institutions preserve discretion over how to respond. They may reward the signal, ignore it, reinterpret it after the fact, or treat it as insufficient. This flexibility has value precisely because it is not paid for. The ability to delay judgment or keep outcomes contingent allows institutions to manage uncertainty without committing their own resources.

Third, institutions may capture reputational rents. The presence of diversity, representation, or symbolic inclusion can generate reputational benefits for firms and organizations, improving external standing or internal legitimacy. These benefits often accrue independently of individual advancement or protection. In transaction-cost terms, this mirrors the logic of appropriable rents described by Klein, Crawford, and Alchian. One party makes a specific investment, while another captures value contingently without guaranteeing a return.

Importantly, this outcome does not require malicious intent. Each institutional decision, such as hiring cautiously, promoting selectively, or adopting a wait-and-see posture, can be locally rational. The asymmetry becomes visible only over time, at the system level rather than

in any single interaction. This explains why the breakdown of cooperation is persistent, widespread, and difficult to contest within standard institutional narratives that focus on isolated decisions rather than cumulative cost burdens.

The losing side in this framework is defined not by identity but by cost position. Diasporic actors bear a range of direct costs in order to sustain cooperation. Financial costs include credential inflation, unpaid internships, relocation expenses, and professional grooming requirements. Cognitive costs arise from constant norm monitoring, self-editing, and anticipatory compliance with poorly articulated expectations. Emotional costs include vigilance, identity suppression, and persistent uncertainty about evaluation criteria. These are real resource expenditures, not expressions of preference or attitude.

A key implication of the model is that exit is delayed rather than immediate. Actors persist because early returns may be positive, cooperation sometimes succeeds, and sunk signaling costs bias continuation. The result is not sudden withdrawal but a slow-drain equilibrium in which continued participation becomes increasingly costly before it becomes untenable.

The framework resolves a central empirical puzzle: the coexistence of individual success stories with persistent group-level breakdown. Signaling works sometimes, reinforcement is inconsistent, and success is stochastic rather than systematic. This produces selection bias in observed outcomes. Institutions observe the survivors rather than the attrited and infer that the system is functioning as intended.

Survivorship narratives further distort diagnosis. When institutions point to individuals who succeed, they treat success as evidence that the system is fair and open. The model shows

instead that success does not falsify asymmetric cost structures. In many cases, it reflects unusually high tolerance for depletion rather than ordinary institutional support. As a result, institutional health is misdiagnosed precisely because costs are unevenly borne.

Because signaling costs are largely unobserved, failure is often attributed to individual deficiencies such as lack of effort, poor fit, cultural mismatch, or weak motivation. These explanations are attractive because they require no institutional adjustment. Responsibility is assigned to the actor who exits rather than to the structure that made continued participation unprofitable.

Polanyi's account of tacit knowledge clarifies why institutions misinterpret their own success. Mature institutions no longer remember how their rules emerged or how costly they once were to learn. Compliance is interpreted as natural rather than as the result of investment. As a result, the invisibility of signaling labor becomes evidence that it was unnecessary. Institutional memory thus works against outsiders by erasing the costs they bear.

When aggregated over time, asymmetric signaling produces systemic effects. Elevated turnover, stalled promotion pipelines, parallel informal institutions, and fragile trust equilibria emerge not through deliberate design but through repeated individual responses to unreciprocated cost burdens. These outcomes are emergent rather than planned.

As cooperation becomes too costly, actors respond rationally by redirecting effort inward, relying on co-ethnic or diasporic networks, and building alternative enforcement mechanisms. This mirrors the logic observed in Greif's account of the Maghribi traders, where high interpretability supported durable cooperation. In the present setting, however, host dominance

and uneven interpretability prevent similar consolidation. Integration stalls not because of hostility but because costs are exhausted.

The first payoff of the framework is an explanation for why cooperation collapses over time even after it begins. Cooperation clears the initial hurdle, but it fails to stabilize because costs are asymmetric, signals are not reciprocated, and governance structures do not realign. This is the combined logic of Williamson and Leeson applied intertemporally rather than statically.

The second payoff is an explanation for why institutions fail in third-culture contexts. Institutions require repeated cooperation, shared meaning, mutually legible signals, and convergent expectations. When one group must continually invest in being understood while the other does not respond in kind, institutions cannot consolidate. Failure is structural rather than moral.

The model does not claim that all cross-cultural cooperation fails, that discriminatory intent is universal, or that outcomes are immutable. It claims only that when the cost of being trustworthy falls asymmetrically on one party, cooperation becomes fragile even among rational, well-intentioned actors. Asymmetric signaling produces quiet institutional erosion rather than visible conflict. The conclusion now asks the final question. If cooperation creates value and conflict destroys it, what would it mean to design institutions in which trust is not privately subsidized by only one side?

VIII. Not a Zero-Sum Game

Cooperation across cultural distance fails not because people are unwilling to cooperate, and not because gains from trade are absent, but because the cost of being trustworthy is distributed asymmetrically. Cooperation is a rule-governed equilibrium. It requires potential

mutual gain, shared rules that are intelligible to participants, and sufficient parity in the capacity to signal, interpret, and enforce those rules. When these conditions hold, cooperation is stable. When they do not, particularly when signaling costs are high, specific, and unobserved, cooperation becomes fragile and erodes over time.

The analysis has shown that diasporic actors respond rationally to social and institutional distance by investing in costly signals that are intended to make cooperation self-enforcing. When those signals are insufficiently legible or reciprocated, actors may escalate commitment further by engaging in what transaction-cost economics would describe as hostage-giving. These hostages take the form of reputational exposure, irreversible self-binding, probationary status, or acceptance of concentrated downside risk. Such strategies can temporarily stabilize cooperation, but they deepen asymmetry by increasing asset specificity and exposure. Rather than resolving the underlying interpretive problem, hostage-giving accelerates depletion when reciprocity does not follow.

The central contribution of this paper is a unified mechanism explaining why cross-cultural cooperation often collapses intertemporally. Diasporic actors invest in costly, specific signals to reduce social and institutional distance, but when host institutions cannot reliably observe, interpret, or reciprocate those signals, the costs of cooperation fall asymmetrically, rendering the cooperative equilibrium unsustainable. This framework synthesizes Williamson's analysis of hazards, asset specificity, and governance misalignment, Leeson's account of norm distance and costly signaling in self-enforcing exchange, the Klein–Crawford–Alchian logic of appropriable rents and hold-up, and the institutional view of culture developed by North, Sowell, and de Soto.

By bringing these literatures together, the paper shows how cooperation can fail without moral failure, without irrationality, and without overt conflict. Breakdown emerges not from refusal to cooperate but from cumulative misalignment between investment, interpretation, and governance.

Nothing in the mechanism implies that cooperation benefits one group only by harming another. Cooperation creates value, conflict destroys it, and misaligned institutions waste surplus that could otherwise be shared. The losses identified here are joint. Marginalized actors lose through depletion, attrition, and exit, while institutions lose through turnover, fragile trust, stalled consolidation, and unrealized gains from trade. The outcome is not redistribution but value destruction through miscoordination.

Even when diasporic actors escalate commitment through hostage-giving, the resulting stability is temporary and inefficient. Concentrating risk on one side may sustain exchange briefly, but it increases fragility rather than resilience. In this sense, asymmetric cooperation is costly for all parties, even when its costs are unevenly borne.

The framework reframes several common debates. First, it shifts attention from prejudice to structure. Persistent disparities need not be explained solely by taste-based or statistical discrimination. They can arise from unobserved signaling asymmetries embedded in institutional design. Second, it redirects analysis from intentions to incentives. Cooperative failure does not require bad actors. It emerges when incentives make trust privately costly for only one side. Third, it reframes inclusion as an interpretive problem. Inclusion without interpretability is unstable. Signals must be not only costly but legible, reciprocated, and protected by governance structures that prevent unilateral exposure.

This paper does not claim that all multicultural institutions fail, that all signaling is unreciprocated, or that all host institutions are incapable of adaptation. It claims only that where cooperation depends on costly, specific signals that one party must continually send and the other does not reliably interpret or reward, the cooperative equilibrium will erode over time. Hostage-giving may delay collapse, but it cannot substitute for shared interpretive capacity or aligned governance.

Cooperation does not collapse because people refuse to cooperate. It collapses because the cost of maintaining trust falls too heavily on one side. If cooperation is to be sustained across cultural distance, institutions must do more than invite participation. They must make trust mutually intelligible, mutually observable, and mutually rewarded. That project is not a charity. It is the institutional precondition for durable growth.

References

Anderson, Elijah. 1999. *Code of the Street: Decency, Violence, and the Moral Life of the Inner City.* New York: W. W. Norton.

Becker, Gary S. 1957. *The Economics of Discrimination.* Chicago: University of Chicago Press.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal?" *American Economic Review* 94 (4): 991–1013.

Buchanan, James M., and Gordon Tullock. 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy.* Ann Arbor: University of Michigan Press.

Butcher, Kristin F., and Anne Morrison Piehl. 2007. "Why Are Immigrant Incarceration Rates Lower? Evidence from California." *Journal of Law and Economics* 50 (2): 459–81.

Coase, Ronald H. 1937. "The Nature of the Firm." *Economica* 4 (16): 386–405.

———. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3: 1–44.

de Soto, Hernando. 2000. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else.* New York: Basic Books.

Farnsworth, Ward. 1999. "Do Parties to Nuisance Cases Bargain after Judgment? A Glimpse inside the Cathedral." *University of Chicago Law Review* 66 (2): 373–436.

Foner, Nancy. 2001. *Islands in the City: West Indian Migration to New York.* Berkeley: University of California Press.

Fryer, Roland G. Jr., and Steven D. Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *Quarterly Journal of Economics* 119 (3): 767–805.

Greif, Avner. 1993. "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition." *American Economic Review* 83 (3): 525–48.

Hayek, F. A. 1960. *The Constitution of Liberty*. Chicago: University of Chicago Press.

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian. 1978. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." *Journal of Law and Economics* 21 (2): 297–326.

Lamb, Vickie Y. M. 1999. "Race, Ethnicity, and the Welfare State: West Indian Immigrants." *Social Service Review* 73 (4): 569–84.

Leeson, Peter T. 2006. "Cooperation and Conflict: Evidence on Self-Enforcing Arrangements and Heterogeneous Groups." *American Journal of Economics and Sociology* 65 (4): 891–907.

———. 2008. “Social Distance and Self-Enforcing Exchange.” *Journal of Legal Studies* 37 (1): 161–88.

Mises, Ludwig von. 1998. *Human Action: A Treatise on Economics.* Auburn, AL: Ludwig von Mises Institute. (Orig. pub. 1949.)

Moss, Philip, and Chris Tilly. 1996. “‘Soft’ Skills and Race: An Investigation of Black Men’s Employment Problems.” *Work and Occupations* 23 (3): 252–276.

North, Douglass C. 1991. “Institutions.” *Journal of Economic Perspectives* 5 (1): 97–112.

Pager, Devah. 2007. *Marked: Race, Crime, and Finding Work in an Era of Mass Incarceration.* Chicago: University of Chicago Press.

Pager, Devah, Bruce Western, and Bart Bonikowski. 2009. “Discrimination in a Low-Wage Labor Market.” *American Sociological Review* 74 (5): 777–99.

Phelps, Edmund S. 1972. “The Statistical Theory of Racism and Sexism.” *American Economic Review* 62 (4): 659–61.

Polanyi, Michael. 1966. *The Tacit Dimension.* Garden City, NY: Doubleday.

Ricardo, David. 1951. *The Works and Correspondence of David Ricardo, Vol. 1.* Edited by Piero Sraffa. Cambridge: Cambridge University Press.

Royster, Deirdre. 2003. *Race and the Invisible Hand: How White Networks Exclude Black Men from Blue-Collar Jobs.* Berkeley: University of California Press.

Smith, Adam. 1759. *The Theory of Moral Sentiments.* Edited by D. D. Raphael and A. L. Macfie. Indianapolis: Liberty Fund, 1982.

Sowell, Thomas. 1981. *Ethnic America.* New York: Basic Books.

Waters, Mary C. 2001. *Black Identities: West Indian Immigrant Dreams and American Realities.* Cambridge, MA: Harvard University Press.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting.* New York: Free Press.

———. 2010. “Transaction Cost Economics: The Natural Progression.” *Journal of Retailing* 86 (3): 215–26.