VOLUME 1 · ISSUE 1 · 14 JULY 2026

The Exodus

A High-School Journal of Law, Economics, and Geopolitics

— THIS ISSUE —

Storming the Bastille

PUBLISHED · 14 JULY 2026

The Exodus

Vol. 1 · No. 1 · Storming the Bastille · July 2026

FOUNDING EDITORIAL BOARD

MOTTO

> *“L’audace, l’audace, encore l’audace, toujours l’audace!”*
> — Georges Jacques Danton

The Exodus is a student-run international journal publishing peer-reviewed essays by high-school researchers on questions in law, economics, and geopolitics. Papers reflect the views of their authors alone. Published semi-annually. Web: ourownexodus.com.

FROM THE EDITORS

A Prison Full of Questions

The Bastille did not fall because it was full. When the crowd came for it on 14 July 1789, the prison held seven inmates: four forgers, an aristocrat, and two men who had lost their minds. What the fortress held was the memory of an order in which such a place made sense — walls high enough that the questions inside them did not have to be answered by anyone standing outside. The crowd came for the memory. They tore down the stones so that the questions could not be locked back up.

This is a journal of undergraduate work, not history. But that image is a good one for a first issue, because the papers collected here all arrive at the same odd doorway from different corridors. Each one finds an old order — a treaty, a market, a border, a rule — that still stands in the shape it was given decades ago, and each one asks, quietly and without theatrics, whether the shape still fits what the world is now asking of it.

Aditya Gupta opens the issue at the rim of a lunar crater, watching a mining crew scoop ice the 1967 Outer Space Treaty never imagined anyone would reach. Anoosha Sharma turns to the 1951 Refugee Convention and the millions whom rising water and failing crops now displace, whom the treaty was not written to see. Armin Vats follows an artist's music into the vaults of the corporation that outlived him. Avni Verma asks what a market is even for, when the thing offered for sale is a kidney or a vote. Anya Varma tracks the quiet redistribution taking place every time a card taps a reader instead of a note changing hands. Anany Srivastava returns to Simon Kuznets and to the reassuring curve that told the twentieth century inequality would fix itself, and finds that it has not. Kabir Srivastava tests the modern instrument of choice — sanctions — against the regimes it was built to move, and finds those regimes have learned. Veer P. Singh sits with a state fragmenting in real time, and the two neighbours pulling at its edges.

There is no single argument uniting these eight papers, and the editorial board has resisted the temptation to invent one. What we will say is that reading them in sequence produced a persistent, uncomfortable feeling: that the frameworks a generation ago called permanent — refugee status, non-appropriation of the commons, price signals as sufficient moral information, sanctions as a lever — are not so much broken as thinly loaded. They were built for a world that no longer exists, and they are being asked to hold a world that is not yet built. The gap between what our institutions say and what our institutions do is widening, and it is widening in exactly the places where the losses are hardest to see: the person without a bank account, the island slipping under the water, the crater not yet claimed.

None of the authors here proposes tearing anything down. They are, if anything, more conservative than the Bastille's crowd — most of them argue for building something modest and workable in the gap, before the gap becomes the norm. That is, we think, the appropriate register for young writers who inherit the world as it is and will spend their lives inside its consequences. It is also, quietly, an act of faith in the old orders themselves: a belief that they can be renegotiated rather than merely worked around.

We are grateful to every contributor for the seriousness of their work, and to the readers now holding this first issue. We hope you find in these pages what we did: eight careful young voices asking the same question in eight different rooms — what still holds, and what has quietly stopped holding, and who is going to be honest enough to say so first.

— The Editorial Board

THIS ISSUE

Contents

PAPER 01 / 08

The Ice Beneath the Loophole

How Moon Mining Quietly Tests a Sixty-Year-Old Promise

Aditya Gupta

ABSTRACT

Start with one stubborn sentence in the 1967 Outer Space Treaty: no country may appropriate the Moon or any other celestial body. For most of the treaty's life that line asked nothing of anyone, because reaching space to take anything was impossible. Mining changes that. This paper works through a single question the treaty never planned for. If the ground cannot be owned, what about the water and metal pulled out of it? Rather than treat that as settled, the paper follows the question through four real responses already on the record: the failed 1979 Moon Agreement, the national mining laws of the United States and Luxembourg, the United States-led Artemis Accords, and the rival lunar partnership of China and Russia. It tests each against the way Earth governs its own shared spaces and lands on an uncomfortable verdict. The working compromise that everyone leans on, own the ore but not the rock, is reasonable and probably unavoidable, yet it carries nothing inside it to stop quiet appropriation of the best sites. The paper closes by sketching a deliberately small international framework, built to be accepted rather than admired, that would let mining proceed without letting it turn into a land grab. *Keywords: space law, Outer Space Treaty, asteroid mining, common heritage of mankind, Artemis Accords, non-appropriation*

The Ice Beneath the Loophole: How Moon Mining Quietly Tests a Sixty-Year-Old Promise Picture a mining crew on the Moon, parked at the rim of a crater that never sees sunlight. Inside it sits ancient ice, the kind you can split into hydrogen and oxygen and pour into a fuel tank. The crew fills a container and lifts off. Have they broken international law? Most people would guess yes, since the Moon is supposed to belong to everyone. The honest answer is that nobody knows, and the reason nobody knows is the heart of this paper. The uncertainty traces back to one line in the 1967 Outer Space Treaty. Space, it says, including the Moon and other bodies, "is not subject to national appropriation by claim of sovereignty, by means of use or occupation, or by any other means" (United Nations, 1967, art. II). For half a century that sentence asked nothing of anybody. You cannot violate a ban on grabbing the Moon when you cannot get to the Moon. Now the getting-there part is arriving. Firms are designing missions to pull water from lunar craters and metals from asteroids, and the moment they do, the sentence stops being a slogan and starts being a problem someone has to solve. The problem has a clean shape. Owning a celestial body is banned outright. Owning what you remove from one is not mentioned at all. Read that gap one way and it is permission. You cannot claim the crater, but the ice in your tank is cargo, the way a fish caught at sea belongs to the boat even though no one owns the water. Read it the other way and it is a loophole wide enough to drive the whole history of conquest through. This paper takes the gap seriously instead of picking a side too early. It asks where the gap came from, watches how four different actors have tried to fill it, borrows what it can from the way Earth runs its shared spaces, and ends with a proposal for closing it before the first serious mining ship clears the pad.

WHAT THE TREATY WAS BUILT TO STOP

A rule makes more sense once you know the fear behind it. The Outer Space Treaty was written in the thick of the Cold War, while two superpowers were hurling rockets past the atmosphere and the rest of the world held its breath. The thing everyone dreaded was not a mining company. It was a sky sliced into national territories and packed with weapons, the rivalry on the ground simply climbing into a place with no way back down. So the treaty did two jobs above all others. It kept the worst weapons out of orbit, and it kept the oldest human habit, claiming new land, off the table. The anti-claiming job was done with unusual force. Article II does not just rule out flags and declarations of sovereignty. It blocks appropriation “by means of use or occupation, or by any other means” (United Nations, 1967, art. II). That last phrase is the drafters covering the exits, trying to stop any clever route to ownership they had not thought of. Two other articles fill in the spirit of the thing. Article I calls space “the province of all mankind” and says its use must serve every country. Article VI makes each government answerable for whatever its private companies get up to in orbit, so no state can launder a land grab through a corporation. Taken together the message is plain. This place stays open to all and closed to capture by any. Notice what the drafters did not write down. There is almost nothing about resources. They were picturing missiles and territorial claims, not a start-up registered in Luxembourg towing a rock toward the Moon. The treaty is generous with lofty principle and stingy with instructions, and that mix tends to crack under pressure. A principle can tell you the Moon belongs to everyone. It cannot tell you whether the crew at the crater rim may keep the ice. It would be a mistake to think the no-claiming rule is just decoration, though. It has bitten before. In 1976 several equatorial countries issued the Bogotá Declaration and tried to claim the

slice of the lucrative geostationary orbit sitting above their territory. The reaction was swift and dismissive. The claim clashed with the treaty, the world said no, and it quietly evaporated. That history matters here. The ban on grabbing space is not a dead letter, which means the mining debate is not about whether you can claim the Moon. You cannot. It is about whether scooping and selling its contents counts as a claim. That is a smaller and harder question, and Bogotá left it wide open.

OWNING A PLACE VERSUS OWNING A THING

Everything turns on a distinction the treaty never spells out. There is a difference between owning a place and owning something taken from that place, and the mining industry has set up camp in exactly that difference. The favorite illustration is the open ocean. No nation owns it. It is the classic example of a thing shared by all and owned by none. Yet a trawler that hauls up a tuna in international waters owns that fish, sells it at port, and nobody accuses the crew of stealing the sea. Taking a resource, the argument goes, is simply a different act from taking the ground beneath it. Lift that reasoning into orbit and the answer seems obvious. The Moon stays unowned while the moonwater becomes property the instant it enters the tank. The reasoning is good enough that it deserves a careful objection rather than a quick dismissal. Two things weaken it. The first is that the ocean is not actually a free-for-all. We fence fishing in with quotas and treaties precisely because letting everyone take whatever they wanted stripped the seas bare, so even at sea the freedom to grab was always policed. The second objection bites harder. Sometimes you cannot separate the resource from the place. A small asteroid might be so modest that mining it means consuming the entire body, which is appropriation of a celestial object by any reading. Park a permanent base on the one reachable

patch of lunar ice, run it for thirty years, and keep all rivals at arm's length, and you are doing something that looks a great deal like occupation, the exact word Article II rules out. The place-versus-thing line is real, but it is a ramp, not a wall, and at the bottom of the ramp taking the ore and taking the ground become the same act. There is a further wrinkle the fishing comparison hides. A tuna is one of millions, and the sea refills. The good spots on the Moon are not like that. There are only so many permanently shadowed craters with reachable ice, and only so many ridges that catch near-constant sunlight for power. These are not renewable stocks but a short list of prime locations, closer to beachfront lots than to fish. When the resource is scarce and fixed in place, owning enough of "the stuff" starts to mean controlling "the place" in every way that matters, even if no flag is ever planted. The analogy that makes mining sound harmless quietly assumes an abundance the Moon does not offer.

THE ANSWER NOBODY ACCEPTED

The world did try to settle this once, head-on, and the attempt is worth studying precisely because it failed. The 1979 Moon Agreement reached straight for the hardest version of an answer. It declared the Moon and its resources "the common heritage of mankind" and called for an international system to manage any future mining, complete with a fair share of the proceeds for every nation (United Nations, 1979, art. 11). Shared ownership, shared control, shared reward. On paper it is the most principled response available. In practice it was a ghost town. A small cluster of states ratified it. Every country actually able to reach the Moon stayed away, the United States, Russia, and China among them. The logic of the refusal is not mysterious. Common heritage means the states doing the expensive and dangerous work must split the winnings with everyone, including nations that risked nothing,

under the eye of some international authority. To the capable, that read as a penalty for ambition. The collapse of the agreement is itself a hard piece of evidence for anything proposed later. The strongest form of shared ownership was laid before the world and the world's rocket-owners walked past it. Any realistic plan has to start from that fact rather than wish it away.

COUNTRIES WRITING THEIR OWN RULES

Where the Moon Agreement left a vacuum, national parliaments rushed in. The United States moved first, with a 2015 law granting any citizen who recovers resources from an asteroid or from space the right to "possess, own, transport, use, and sell" them (Commercial Space Launch Competitiveness Act, 2015). Read it closely and the caution stands out. It never claims sovereignty over any celestial body. It never says America owns the Moon. It hands private operators the ore while loudly declining the rock, threading the exact needle Article II leaves. Luxembourg followed in 2017 with a blunter, shorter law declaring that space resources “are capable of being owned,” wrapped in a tax-and-licensing package designed to lure mining firms to set up there (Grand Duchy of Luxembourg, 2017). The United Arab Emirates and Japan soon passed laws in the same spirit. The pattern is now clear. Instead of waiting on a global deal that may never land, states are manufacturing certainty at home and betting that ownership backed by a serious government is enough to satisfy investors. A question that was supposed to belong to all of humanity is being answered quietly, one national statute at a time. Is that legal? Genuinely unsettled. Defenders say these laws sit comfortably inside the treaty, because they govern resources rather than territory and because Article VI plainly expects states to license and supervise private activity. Critics counter that you cannot give away a right you do not hold. If no one may appropriate the Moon, no parliament can carve off a share of the Moon’s contents and stamp it private property, because the thing being divided was never that

country’s to divide. The fight does not resolve, and it will not, because the treaty everyone keeps citing simply has nothing to say on the point. Each camp is hearing its own answer echo back out of the silence. A second force runs under all this lawmaking, and it is plain competition. Once one country offers mining ventures legal certainty, the rest feel the pull to match it or watch the firms, and the jobs and revenue and prestige they carry, register somewhere friendlier. Luxembourg said out loud that it wanted to be Europe’s hub for the trade, and its success bred copies. This is the familiar race to court a footloose industry, only now the prize on the table is the contents of the Moon. The trouble with such races is direction. They push standards toward whatever pleases the industry being chased, which is the worst possible compass for a question meant to be answered in everyone’s interest. Run long enough, the race does not discover the fair rule. It discovers the loosest one. There is also a problem waiting a little further down this road. National property rights work on Earth because, in the end, courts and governments recognise each other’s claims. What happens when an American company and a Chinese one both want the same crater, each holding a licence from its own government and neither recognising the other’s? There is no shared registry, no common court, and no agreed first-come rule to break the tie. Each operator is legally secure at home and legally invisible to its rival. Domestic certainty, in other words, buys far less than it seems to, because the one place a dispute would actually arise is the one place no national law reaches. A patchwork of confident statutes that contradict one another is not the same thing as a settled question. It is a collision waiting for its moment.

SOFT RULES, SAFETY ZONES, AND TWO CAMPS

Aware that a pile of national laws is no substitute for shared ones, the United States launched the Artemis Accords in 2020, a set of non-binding principles for lunar and asteroid work that other states may sign (National Aeronautics and Space Administration, 2020). The Accords assert that resource extraction can be done in line with the Outer

Space Treaty, which answers the open question by simply announcing the answer. They also introduce a clever device for keeping operators out of each other's way, the "safety zone," a buffer around an activity where everyone else is asked to coordinate before coming close. The safety zone is where the cleverness curdles. Its stated purpose is fair. Two crews working the same patch of regolith need a way to avoid collisions and flying debris. But a buffer you can keep others out of, held for as long as your operation runs, looks a lot like the occupation Article II forbids. Call it safety and it is coordination. Describe what it does and it is a perimeter. The Accords swear the zones are not claims of sovereignty, and as a statement of intention that is surely true. The danger has nothing to do with anyone's motives. It is structural. Stack up enough long-lived exclusive zones, one mission at a time, and the best real estate, the sunlit ridges, the easy ice, ends up under the lasting control of whoever arrived first, and the forbidden word never once gets spoken. Then the politics arrive. Artemis is an American initiative, and while many states have signed on, two heavy powers have pointedly not. Russia and China have built their own partnership around an International Lunar Research Station. What we are watching is the early outline of two camps, each writing its own understanding of what is permitted, each able to claim loyalty to a treaty loose enough to shelter both readings at once. This is how soft rules can set the wrong way. If the major spacefaring states drift, through parallel habit rather than any shared

text, toward rules that tolerate quiet control of the prime sites, that drift can become the working law of the frontier long before a treaty ever blesses it, and it can lock in the first arrivals for good.

THREE LESSONS FROM EARTH'S SHARED SPACES

Humanity has governed unowned places before, and three of those experiments throw useful light, each in a different direction. The open ocean teaches a double lesson. It proves a place can be shared while the resources hauled from it are owned, since the freedom to fish coexists with the rule that no one owns the sea. But it also shows that an unpoliced freedom to take ends in collapse and quarrel, which is why that freedom now sits inside quotas and treaties. The catch-belongs-to-the-boat idea works, in other words, but only inside a fence of rules the romantic early version never had. Antarctica points the other way, toward patience. Rather than resolve the rival territorial claims, the 1959 Antarctic Treaty froze them, set the continent aside for science and peace, and later banned mineral mining outright (Antarctic Treaty, 1959). Sometimes the wisest move is to take the explosive question off the table and let cooperation grow habits of its own. The weakness is obvious. A freeze holds only while the buried resources are not worth enough to shatter it, and lunar fuel may prove far more tempting than anything locked under Antarctic ice. The deep seabed is the closest match of all, and the most sobering. The Law of the Sea Convention names it common heritage and places it under an International Seabed Authority that licenses mining and is meant to share the gains with everyone (United Nations, 1982, pt. XI). This is the very model the Moon Agreement reached for. Its record warns against copying it wholesale. The system is real and it works, but it is slow, contested, and resented by some industrial states as a brake on enterprise, and the United States has never ratified the Convention

partly for that reason. So the seabed proves two things at once. An international licensing-and-sharing body can be built. And building one that satisfies both the capable and the excluded is punishingly hard.

FOLLOW THE MONEY, THEN THE POWER

None of this is really a tidy legal puzzle, because the law is being pushed from behind by money and pulled from the front by power. Strip away the treaty language and two plain motives explain most of the action. The first is the savage cost of leaving Earth. Lifting a single kilogram off the planet costs a fortune, so water already sitting on the Moon is worth far more up there than the same water could ever be down here, because you can crack it into fuel without ever paying to haul it up. The headline prize is not really platinum. It is the chance to refuel ships in space and stop dragging their fuel out of Earth's gravity. Control the reachable lunar ice and you control the filling stations of the solar system. The second motive is scarcity at home. Many of the metals packed into asteroids are the very materials nations already scrap over on Earth, where supply clusters in a few countries and gets treated as a security matter. A state that secures an off-world source gains leverage, not just profit, and a state afraid of being shut out has every reason to stake an early claim in law and on the surface. That is why space resources cannot be quarantined as a commercial question. They rest on top of the same fight over critical materials that already shapes trade and alliances, and they haul that fight up into orbit. Which is what makes the two-camp pattern dangerous. When the leading space power gathers a coalition around its reading of the treaty, and a rival gathers a counter-coalition around a different one, the disagreement stops being academic and becomes two blocs laying down facts on the surface while the law watches. The real risk is not mainly a weapon going off. It is slower

and quieter. Sites get occupied, exclusive zones become ordinary, and a generation of practice piles up that answers the ownership question in favor of the strong and the early, leaving the treaty's fine words upright but hollow. Practice hardens into law, and that law tends to bless whoever moved first. For the many countries without a rocket, the nightmare is not a dramatic treaty breach. It is a slow, deniable erosion that hands over the commons before the rules are ever written.

A SMALL FRAMEWORK, BUILT TO BE SIGNED

Set the pieces side by side and the trajectory is easy to read. The strongest shared-ownership option was offered in 1979 and refused. The capable states have since settled, through national laws and soft accords, on a setup that permits private taking and tolerates exclusive zones, all while pledging fidelity to a treaty that bans appropriation. A rival camp is forming next to them. Nothing binding ties any of it together. Leave it alone and the frontier ends up run by whoever arrives first and can hold what they take, appropriation in fact and denied in name. My argument is that the working compromise, own the ore and not the rock, is the right starting point and the wrong stopping point. Right to start there, because the alternative, a strict common-heritage regime, has already been offered to the exact states whose participation it needs, and they walked away. Demanding it again is a way to be correct in theory and ignored in practice. Wrong to stop there, because once you hand the compromise to national laws and rival blocs, it carries nothing inside it to slow its own slide into real, unspoken appropriation. The catch-belongs-to-the-boat principle worked at sea only after rules were wrapped around it. In space it is being switched on with no wrapping at all.

So the sensible target is a middle path, an international framework lighter than the seabed authority but firmer than today's scatter, ideally grown through the United Nations bodies that already handle space, where lawyers have in fact begun drafting principles for resource work. It would not need a heavy benefit-sharing tax that scares the capable off again. It could aim narrowly at the points where the unregulated version turns dangerous. Cap the size and lifespan of exclusive zones so that staying clear of a neighbor never becomes holding the ground forever. Require operators to register and disclose what they are doing, so the practice stays visible. And open at least a modest channel through which countries without space programs share in the gains from a place the treaty still calls the province of all mankind. The goal is not to win 1979 on a second try. It is to give the own-the-ore compromise the surrounding rules the oceans proved you cannot skip. Timing is half the argument. A deal like this is far easier to strike before the facts harden than after, because once a state or a company holds a valuable site it has every reason to defend the practice that got it there and none to accept new limits. The window is open now precisely because nobody has mined at scale, so there are almost no entrenched positions to protect. It will not stay open. Each year of unregulated activity converts a question that talking could settle into one that can only be settled by going along with what the first movers already did. The case for moving now is not idealism. It is a sober read of how resource regimes set like concrete.

THE QUESTION EVERYONE SKIPS: WHO ENFORCES IT?

Most of this debate argues about what the rule should say. Far less of it asks who would make anyone follow the rule, and on the Moon that second question may matter more. International law has no police force. It runs on reciprocity, reputation, and the willingness of states to call each other out, and all three weaken badly once the action moves a few hundred

thousand kilometres away, where almost nobody can see what is happening and even fewer can do anything about it. A framework that looks strong on paper means little if the only witnesses to a violation are the company committing it and the government that licensed it. This is why the dull-sounding parts of my proposal, registration and disclosure, are doing more work than they appear to. The first problem on the Moon is not punishment. It is simply knowing. If operators must register their activities and report where they are working and how large a zone they are holding, then a quiet expansion of an exclusive area becomes visible, and a visible violation is one that allies and rivals can at least name. Transparency does not enforce a rule by itself. What it does is make enforcement possible, by turning a private act in an unwatched place into a public fact the rest of the world can respond to. Verification, not punishment, is the realistic first line of defence, and any framework that ignores it is writing rules for a room with the lights off.

THREE FAIR OBJECTIONS

The strongest objection is about investment. Any framework, however light, adds cost and doubt, and the fastest way to get humanity mining is to let the bold go first and tidy up the law later. There is real force here. But tidy up later is exactly how first movers cement advantages that everyone after them simply inherits, and Earth's own resource rushes are not a cheerful guide to letting the spoils fall to whoever shows up with the biggest machines. Certainty is itself an asset, and a rule everyone has accepted outlasts a right granted by one parliament that a rival can refuse to recognize. A second objection runs from sovereignty. The capable states will never put their space industries under outside control, and the fate of both the Moon Agreement and, for the United States, the Law of the Sea Convention seems to prove it. The point is serious, and it is the reason

the framework here is kept small on purpose, asking for transparency and limits on exclusion rather than a powerful authority sitting over everyone. The lesson of past refusals is not that international rules are hopeless. It is that maximal ones are. The craft lies in finding the most the capable will sign, not the most the excluded would wish for. The last objection comes from the opposite shore. A light-touch system, it says, sells out the common-heritage ideal, blessing private extraction of what was meant for everyone and telling the rest of the world to make do with transparency and crumbs while a few players grow rich off the sky. The objection carries real moral weight and cannot be fully answered, only weighed. The reply is that the strong version is not on offer. It was, in 1979, and the spacefaring powers refused it. A modest set of shared rules that states actually obey guards the common interest better than a magnificent principle they all ignore. Half a loaf is not justice. But the real choice on this frontier is not half a loaf against a whole one. It is half a loaf against a free-for-all.

CONCLUSION

Go back to the crew at the crater rim, container full, lifting off. The treaty written in 1967 cannot tell them whether what they just did was lawful, because the people who wrote it were guarding against flags and missiles, not drills and fuel tanks. They forbade the conquest they could imagine and stayed silent on the commerce they could not. Into that silence the capable states have stepped, answering the question for themselves, one statute and one accord at a time. The honest summary is that the famous ban on appropriation is not being repealed. It is being walked around. The words stay sacred while the practice drifts beneath them, until the distance between what the treaty says and what operators do grows wide enough that the words no longer match the world. That distance can be closed two ways. By drift, as national laws and rival blocs settle through habit and first arrival into a regime of possession the text politely refuses to admit. Or by design, through a small, negotiated framework that accepts the mining the capable states will do regardless and rings it with the limits, the openness, and the modest sharing that keep mining from sliding into conquest. No one owns the Moon. The law still says so, and that is worth keeping true and not merely keeping written. Whether it stays true depends less on how gracefully a sentence from 1967 reads and more on whether the spacefaring powers can manage, this time, the cooperation an earlier generation found when the stakes were missiles instead of minerals. The crew is almost ready to fly. The law should not wait for them to land.

REFERENCES

Antarctic Treaty, December 1, 1959, 402 U.N.T.S. 71. Commercial Space Launch Competitiveness Act, Pub. L. No. 114-90, 129 Stat. 704 (2015). Grand Duchy of Luxembourg. (2017). Law of 20 July 2017 on the exploration and use of space resources. Lee, R. J. (2012). Law and regulation of commercial mining of minerals in outer space. Springer. National Aeronautics and Space Administration. (2020). The Artemis Accords: Principles for cooperation in the civil exploration and use of the Moon, Mars, comets, and asteroids for peaceful purposes. United Nations. (1967). Treaty on principles governing the activities of states in the exploration and use of outer space, including the Moon and other celestial bodies (United Nations Treaty Series, Vol. 610, p. 205). United Nations. (1979). Agreement governing the activities of states on the Moon and other celestial bodies (United Nations Treaty Series, Vol. 1363, p. 3). United Nations. (1982). United Nations Convention on the Law of the Sea (United Nations Treaty Series, Vol. 1833, p. 397). von der Dunk, F. (Ed.). (2015). Handbook of space law. Edward Elgar.

PAPER 02 / 08

Stateless by Degrees

The Legal Lacuna Facing Climate-Displaced Persons Under International Refugee Law

Anoosha Sharma

ABSTRACT

The displacement of people by environmental change is among the defining humanitarian challenges of the twenty-first century, yet the principal instrument of international protection---the 1951 Convention Relating to the Status of Refugees---offers those displaced little to no formal coverage. This article examines why. It argues that the gap is not an oversight to be patched but a structural feature of a treaty built around state persecution and individualized fear. Drawing on the text of the Convention, the landmark decision in Teitiota v. New Zealand, and the patchwork of regional and soft-law instruments that has emerged in its shadow, the article evaluates four competing reform pathways: a dedicated climate-displacement treaty, reinterpretation of the existing Convention, regional and bilateral mobility frameworks, and the extension of human-rights-based non-refoulement. It concludes that no single instrument can close the lacuna, and that the most defensible path is a layered regime: a non-refoulement floor grounded in human rights law, regional mobility agreements tailored to disaster realities, and planned relocation pathways for territorially threatened states. The analysis cautions that legal categorization, while necessary, cannot substitute for the political will to share responsibility across borders.

KEYWORDS *climate displacement, refugee law, non-refoulement, statelessness, Teitiota, planned relocation*

In 2013, Ioane Teitiota, a national of the low-lying Pacific state of Kiribati, asked New Zealand for protection. His home was not threatened by an army or a government but by the sea: saltwater intrusion had contaminated the freshwater supply, erosion was consuming habitable land, and the cumulative effects of rising sea levels had, in his account, rendered ordinary life increasingly untenable. New Zealand refused his claim, and its courts upheld the refusal on a deceptively simple ground. Teitiota was not, in the legal sense the word carries under international law, a refugee. He feared no persecutor. The harm he fled was indifferent to his race, religion, nationality, political opinion, or social group. He was, in the eyes of the governing treaty, a migrant fleeing misfortune rather than a refugee fleeing persecution.

That distinction---between misfortune and persecution---sits at the center of one of the most consequential gaps in contemporary international law. Estimates of the number of people who may be displaced by environmental change in the coming decades vary widely and are notoriously contested, but even conservative projections describe movement on a scale that existing legal categories were never designed to absorb. The vocabulary that has grown up around the phenomenon---"climate refugees," "environmental refugees"---reflects an intuition that these people deserve protection comparable to those fleeing war or oppression. Yet the vocabulary is legally inaccurate, and the inaccuracy is not merely pedantic. It points to a substantive truth: the architecture of international refugee protection does not, and on its current terms cannot, accommodate those displaced primarily by environmental change.

This article asks why that is so, and what ought to be done about it. It proceeds in five parts. The first reconstructs the legal definition of a refugee under the 1951 Convention and its 1967 Protocol, identifying the elements that a claimant must satisfy. The second explains why climate-displaced persons systematically fail to satisfy those elements, and why the failure is structural rather than incidental. The third turns to jurisprudence, analyzing the 2020 decision of the United Nations Human Rights Committee in Teitiota v. New Zealand, which---while rejecting the individual claim---articulated a principle with potentially far-reaching consequences. The fourth surveys the regional instruments and soft-law initiatives that have emerged to fill the vacuum. The fifth evaluates the competing reform proposals and defends a particular synthesis. The central argument is that the search for a single, elegant legal solution is misguided: the protection gap will be narrowed, if at all, by a layered regime combining a human-rights-based duty of non-refoulement, regionally negotiated mobility frameworks, and anticipatory pathways for the inhabitants of states facing the loss of territory itself.

THE ARCHITECTURE OF THE REFUGEE DEFINITION

The 1951 Convention Relating to the Status of Refugees, as amended by its 1967 Protocol, defines a refugee as a person who, "owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion," is outside the country of nationality and is unable or unwilling to avail themselves of that country's protection (United Nations, 1951, art. 1A\[2\]). Embedded in this single sentence are several distinct requirements, each of which functions as a filter. A claimant must satisfy every one; failure on any single element defeats the claim.

The first requirement is a well-founded fear, a standard understood to contain both a subjective element (the claimant genuinely fears harm) and an objective element (the fear is reasonable in light of conditions in the country of origin). The second is that the feared harm rise to the level of persecution. Although the Convention does not define persecution exhaustively, it is generally understood to involve serious violations of human rights, frequently though not exclusively threats to life or freedom. The third, and for present purposes the most decisive, is the nexus requirement: the persecution must be inflicted for reasons of one of five enumerated grounds. The harm must be tethered to the claimant's race, religion, nationality, political opinion, or membership in a particular social group. A person who fears serious harm for reasons falling outside these five grounds is not, however grave the harm, a Convention refugee.

The fourth requirement is alienage: the claimant must have crossed an international border and be outside the country of nationality. A person who has fled their home but remains within their own state's territory is not a refugee but an internally displaced person, a category that the Convention does not reach at all. The fifth is the failure of state protection: the country of origin must be unable or unwilling to protect the claimant from the feared harm. Underlying the entire structure is an implicit model of the refugee as a person fleeing their own government, or fleeing harm that their own government cannot or will not prevent. The refugee, in the Convention's imagination, is a citizen whose bond of protection with the state has broken down.

This model was a product of its moment. Drafted in the aftermath of the Second World War and shaped by the experience of European populations displaced by totalitarian persecution, the Convention codified a particular and historically specific understanding of forced migration: flight from a hostile or failed state. Whatever its merits for that paradigm, the model maps poorly onto displacement driven by environmental change, as the next section explains.

WHY CLIMATE DISPLACEMENT FALLS OUTSIDE THE FRAMEWORK

The mismatch between climate displacement and the refugee definition is not the result of a single missing element but of several, each independently sufficient to defeat a claim. Understanding them in turn clarifies why incremental fixes are so difficult.

The Absence of a Persecutor

Persecution, as the Convention conceives it, is something done to a person by an agent. The paradigmatic persecutor is a state; in some interpretations, non-state actors whom the state cannot control may qualify. But rising seas, desertification, and intensifying storms are not agents. They do not single out their victims for reasons of identity or belief. A person whose island is becoming uninhabitable cannot point to a persecutor who intends them harm. The harm is real and may be lethal, but it is, in the law's terms, undirected. This is the deepest of the structural problems, because persecution is the conceptual heart of the refugee definition, and environmental harm is categorically unlike it.

The Nexus Problem

Even if environmental harm could somehow be characterized as persecution, the claimant would still face the nexus requirement: the harm must be inflicted for reasons of one of the five enumerated grounds. Climate change does not discriminate on the basis of race, religion, nationality, political opinion, or social group. It is true that its effects fall unevenly, and that the poorest and most marginalized populations frequently bear the heaviest burdens. Scholars and advocates have sometimes attempted to reframe climate displacement through this lens, arguing that where environmental harm compounds existing patterns of discrimination, a nexus to a protected ground may be established. Such arguments may succeed at the margins, in cases where a government deliberately withholds disaster relief from a disfavored group, for instance. But they cannot convert the generalized phenomenon of climate displacement into a Convention claim, because the underlying environmental harm is not itself imposed for a protected reason.

The Internal Displacement Majority

A further difficulty is empirical. The Convention protects only those who have crossed an international border, yet the great majority of people displaced by environmental events do not leave their own countries. They move from coast to interior, from countryside to city, remaining within national territory. These internally displaced persons fall entirely outside the Convention regardless of the severity of their plight. The cross-border cases that do reach the asylum system are thus the visible minority of a much larger phenomenon, and the legal category that governs them was designed for a different kind of border crossing altogether.

Slow Onset, Sudden Onset, and the Problem of Causation

Environmental displacement resists the law's preference for discrete, identifiable events. Sudden-onset disasters---a cyclone, a flood---produce displacement that is at least temporally legible, even if the Convention does not cover it. Slow-onset processes---sea-level rise, salinization, the gradual failure of agriculture---unfold over years or decades and interact with economic, political, and demographic pressures in ways that make it nearly impossible to isolate "the environment" as the sole or even primary cause of any given person's movement. This causal entanglement is not merely an evidentiary inconvenience. It reflects the genuine multiplicity of forces that drive migration, and it frustrates any attempt to define a clean legal category of persons displaced "by" climate change.

The Special Case of Disappearing States

The most extreme scenario presents a problem the Convention does not even contemplate: the physical disappearance of a state. For low-lying island nations such as Tuvalu, Kiribati, and the Marshall Islands, sustained sea-level rise raises the prospect that national territory may become uninhabitable or vanish altogether. If a state ceases to have a habitable territory and a resident population, the consequences for its nationals extend beyond displacement to the risk of statelessness---the loss not merely of a home but of a nationality and the protections that flow from it. The refugee framework, which presupposes a functioning state of origin from which the claimant has fled, is conceptually unequipped to address a situation in which the state of origin itself has ceased to exist in any meaningful sense. This scenario, though still prospective, exposes the limits of every existing category and has begun to drive some of the most innovative legal responses, discussed below.

THE NON-REFOULEMENT TURN: TEITIOTA V. NEW ZEALAND

If the Refugee Convention is a closed door, the more promising opening has come from a different body of law. In 2020, the United Nations Human Rights Committee delivered its views in *Teitiota v. New Zealand*, the communication arising from the case described at the outset of this article (Human Rights Committee, 2020). The Committee's reasoning repays close attention, because it simultaneously rejected the individual claim and articulated a principle that may prove more significant than the outcome.

The Committee considered the claim not under the Refugee Convention but under the International Covenant on Civil and Political Rights, specifically the right to life protected by Article 6. The relevant question was whether New Zealand, by returning Teitiota to Kiribati, would expose him to a real risk of irreparable harm to his right to life, thereby violating the principle of non-refoulement as it operates under the Covenant. Non-refoulement---the prohibition on returning a person to a place where they face certain grave harms---is the cornerstone obligation of international protection, and it is not confined to the Refugee Convention. It binds states under human rights treaties independently of refugee status.

On the facts, the Committee found no violation. It accepted New Zealand's assessment that the risks Teitiota faced were not sufficiently imminent: the timeframe over which Kiribati might become uninhabitable---a matter of years or decades---left room for intervening measures by the state and the international community, and Teitiota had not established that he faced an immediate risk to his life. The individual claim therefore failed. But the Committee went further than the result required. It recognized, as a matter of principle, that environmental degradation and the effects of climate change can expose individuals to violations of the right to life, and that where such risks are sufficiently serious, the principle of non-refoulement may prohibit a state from returning a person to those conditions. The Committee observed, in effect, that the international community's obligations could be engaged before a country becomes uninhabitable, and that states must not deport individuals into conditions that would imperil their lives.

The significance of *Teitiota* lies in this dictum. For the first time, an international human rights body acknowledged that climate-related harm could, in principle, trigger non-refoulement obligations. The decision did not create a new protected status, and it set the evidentiary threshold high, requiring something approaching an imminent threat to life. But it relocated the question of climate displacement from refugee law, where it is structurally barred, to human rights law, where the governing concept---risk to life---is indifferent to whether the source of the risk is a persecutor or a rising sea. In doing so it suggested a route by which protection might be extended without amending the Refugee Convention at all.

The limits of the route should not be understated. A non-refoulement claim grounded in the right to life requires the claimant first to have reached the territory or jurisdiction of a protecting state, and then to demonstrate a risk of a severity that slow-onset environmental change will rarely satisfy until it is far advanced. It offers nothing to the internally displaced, and little to those whose circumstances, though dire, fall short of an imminent threat to life. It is a floor, not a comprehensive regime. But a floor grounded in binding human rights obligations is a more secure foundation than the aspirational language of soft law, and any durable architecture of protection is likely to build upon it.

THE PATCHWORK: REGIONAL AND SOFT-LAW RESPONSES

In the absence of a global treaty, a fragmented body of regional instruments and non-binding initiatives has developed. None closes the gap, but together they sketch the elements from which a more coherent regime might eventually be assembled.

Regional Instruments

Two regional refugee instruments adopt definitions broader than the 1951 Convention. The 1969 Convention of the Organisation of African Unity governing refugee problems in Africa, and the 1984 Cartagena Declaration on Refugees adopted by Latin American states, both extend protection to persons fleeing events seriously disturbing public order. While neither was drafted with environmental displacement in mind, their broader language has been read by some commentators as capable of encompassing certain disaster-driven movements. More directly relevant is the African Union's 2009 Kampala Convention, the first binding regional treaty to address internal displacement comprehensively, which expressly includes persons displaced by natural or human-made disasters, including the effects of climate change (African Union, 2009). The Kampala Convention is notable precisely because it reaches the internally displaced, the population the global refugee regime ignores. Its limitation is geographic: it binds only African Union member states and addresses only internal displacement.

Soft-Law Initiatives

At the global level, the response has been confined largely to non-binding instruments. The Nansen Initiative, a state-led process launched in 2012, produced in 2015 a Protection Agenda addressing cross-border displacement in the context of disasters and climate change; it was succeeded by the Platform on Disaster Displacement, which continues to promote the Agenda's recommendations. The 2018 Global Compact for Safe, Orderly and Regular Migration explicitly recognizes environmental degradation and climate change as drivers of migration and calls on states to develop adaptation and mobility responses (United Nations, 2018). These instruments have value: they build consensus, develop vocabulary, and create soft expectations that can harden over time. But they are, by design, non-binding. They impose no enforceable obligation on any state to admit or protect a single displaced person.

Bilateral Innovation: The Falepili Union

The most striking recent development is bilateral. In 2023, Australia and Tuvalu concluded the Falepili Union treaty, which---alongside security provisions---establishes a special mobility pathway permitting a capped number of Tuvaluan citizens to migrate to Australia each year, framed explicitly around the existential threat that climate change poses to Tuvalu (Government of Australia & Government of Tuvalu, 2023). The arrangement is significant as the first treaty to create a dedicated migration pathway in anticipation of climate-driven displacement, and as a model of how planned, orderly mobility might be negotiated before crisis forces disorderly movement. Its limits are equally instructive: it is bilateral, capped, and accompanied by provisions giving Australia a degree of influence over Tuvalu's security arrangements, prompting debate about the price of such agreements for the sovereignty of vulnerable states.

COMPETING REFORM PATHWAYS

Four broad strategies dominate the debate over how to close the protection gap. Each has prominent advocates; each carries characteristic risks.

A Dedicated Climate-Displacement Treaty

The most intuitive proposal is a new global treaty creating a recognized status for climate-displaced persons, with corresponding obligations of admission and protection. Its appeal is clarity: a purpose-built instrument could define the protected class, allocate responsibility among states, and establish institutional machinery. Its weakness is political. The Refugee Convention itself would almost certainly not be adopted if proposed today, and the appetite of states for new obligations to admit displaced foreigners is, if anything, diminishing. A negotiation begun in the present climate of restrictionism would risk producing either no agreement or an agreement so weakened by exceptions as to offer less than existing human rights law already provides. There is also a subtler danger: reopening the definition of who deserves protection could invite attempts to narrow the existing refugee definition rather than supplement it.

Reinterpreting the Existing Convention

A second strategy would stretch the 1951 Convention through interpretation, building on the marginal arguments noted earlier to bring at least some climate-displaced persons within its terms. This approach has the advantage of requiring no new treaty and no fresh consent from states. But it strains against the text. The persecution and nexus requirements are not interpretive details that can be argued away; they are the load-bearing elements of the definition. An interpretation expansive enough to cover generalized climate displacement would be difficult to reconcile with the Convention's language and would likely be resisted by states precisely because of its open-endedness. Interpretation can do useful work at the edges, but it cannot bear the weight of the whole problem.

Regional and Bilateral Mobility Frameworks

A third strategy abandons the search for a global instrument in favor of regional and bilateral arrangements tailored to particular geographies and relationships. The Kampala Convention and the Falepili Union exemplify the approach. Its strength is feasibility: agreements among states with shared interests and existing ties are easier to negotiate and more likely to be honored than universal commitments. Its weakness is coverage. A patchwork of regional and bilateral deals will inevitably leave gaps, will treat similarly situated people differently depending on geography, and will depend on the goodwill of wealthier states toward particular neighbors. It distributes protection by accident of location rather than by principle.

Human-Rights-Based Complementary Protection

The fourth strategy builds on *Teitiota*: rather than create a new status, it extends protection through the non-refoulement obligations already binding on states under human rights law, particularly the right to life and the prohibition on cruel, inhuman, or degrading treatment. This is the mechanism of "complementary protection," under which a person who does not qualify as a refugee may nonetheless not be returned to a place of grave harm. Its great advantage is that it rests on obligations states have already accepted, requiring no new consent. Its limitation, as noted, is its high threshold and its inapplicability to the internally displaced and to those whose circumstances fall short of an imminent threat to fundamental rights.

ANALYSIS: THE CASE FOR A LAYERED REGIME

The temptation, confronting a gap of this magnitude, is to seek a single instrument that fills it. The argument of this article is that the temptation should be resisted, for two reasons. The first is diagnostic: the phenomenon of climate displacement is not one problem but several. It encompasses the internally displaced and the cross-border migrant, the victim of the sudden cyclone and the farmer slowly defeated by salinization, the citizen of a stressed state and the national of a state facing physical disappearance. No single legal category can fit populations this heterogeneous without either excluding most of them or becoming so capacious as to be meaningless. The second reason is prudential: in the current political environment, the negotiation of an ambitious new global treaty is more likely to produce a weak instrument, or none, than a strong one, and the attempt could destabilize the protections that already exist.

A more defensible architecture is layered. Its foundation is the non-refoulement floor identified in *Teitiota*: a binding obligation, grounded in human rights law and applicable to all states party to the relevant covenants, not to return individuals to conditions posing a real risk to life or to freedom from inhuman treatment, whatever the source of that risk. This floor requires no new treaty and is already, in principle, the law. The task is to develop it---through litigation, through the considered views of treaty bodies, and through state practice---into a more predictable and lower-threshold protection than *Teitiota*'s demanding imminence standard currently allows.

Above this floor sits the regional and bilateral layer: agreements such as the Kampala Convention and the Falepili Union, which can address the internally displaced and create orderly mobility pathways that human rights law alone does not provide. The deficiency of this layer---its unevenness---is real but can be mitigated if regional arrangements are encouraged to converge toward common standards, perhaps guided by the soft-law frameworks already developed through the Nansen Initiative and the Global Compact. Soft law, on this view, is not a failure but a function: it is the laboratory in which the norms that may one day harden are developed and tested.

The final layer addresses the hardest case, the disappearing state. Here the appropriate response is anticipatory rather than reactive. Planned relocation negotiated in advance, the preservation of nationality and statehood through legal arrangements that do not depend on continued habitability of territory, and dedicated migration pathways of the Falepili type are the tools available. The defining virtue of anticipation is that it converts what would otherwise be a sudden humanitarian emergency, met by an overwhelmed and reluctant asylum system, into an orderly process negotiated while time remains. The decision in Teitiota itself gestured toward this logic when it observed that the years before a state becomes uninhabitable are years in which protective action remains possible.

The layered regime is admittedly less satisfying than a single grand instrument. It offers no clean category, no universally recognized status, no tidy allocation of responsibility. But it has the cardinal virtue of building on foundations that already exist and on consent that states have already given, rather than staking the protection of millions on a negotiation that the present moment is unlikely to deliver. It treats the protection gap not as a hole to be filled by one act of legislation but as a set of related deficiencies to be narrowed by complementary means.

COUNTERARGUMENTS AND LIMITATIONS

Three objections to this argument deserve direct acknowledgment. The first is that a layered, largely human-rights-based regime underprotects: it offers a high-threshold floor and a patchwork above it, leaving many displaced persons---especially the internally displaced and those facing slow-onset harm short of an imminent threat to life---with no enforceable protection at all. This is true, and it is the strongest objection. The response is not that the layered regime is adequate but that it is more attainable than the alternative, and that an attainable partial protection is preferable to an unattainable comprehensive one. The argument is comparative, not complacent.

The second objection runs from the opposite direction: that even the modest extension of non-refoulement contemplated here will provoke a sovereignty backlash, as states resist any obligation that constrains their control over admission. This concern is realistic, and it is precisely why the article does not advocate a new treaty: the layered approach asks states for little they have not already granted, relying on the development of existing obligations rather than the creation of new ones. The risk of backlash is lowest where the legal foundation is most firmly established, which is an argument for building on binding human rights commitments rather than negotiating fresh ones.

The third objection is normative. A regime that distributes protection through regional and bilateral arrangements, and that conditions the most generous pathways on the goodwill of wealthy states, entrenches inequity: it makes the protection a displaced person receives depend on where they happen to live and on the strategic interest a powerful neighbor takes in their fate. The Falepili Union, with its accompanying constraints on Tuvaluan autonomy, illustrates the danger. This objection cannot be fully answered within the framework of the layered regime; it can only be mitigated by insisting that regional arrangements converge toward common minimum standards and that the bargains struck respect the sovereignty of the vulnerable states that enter them. The objection is a reminder that legal architecture is not a substitute for justice in the distribution of responsibility, a point the conclusion takes up.

BEYOND LAW: THE ECONOMIC AND GEOPOLITICAL DIMENSIONS

Because climate displacement sits at the intersection of law, economics, and geopolitics, a purely doctrinal analysis is incomplete. The legal categories examined above are downstream of a distributive question that is fundamentally economic: who bears the cost of displacement, and on what principle. The populations most exposed to climate-driven displacement are concentrated in low-income and lower-middle-income states that have contributed least, in historical terms, to the accumulation of greenhouse gases. The wealthier states whose emissions drove much of the warming are, by and large, geographically and economically better insulated from its displacement effects. This asymmetry between contribution and exposure is the moral engine of the entire debate, and it explains why arguments framed solely around the receiving state's discretion feel inadequate: they treat as a matter of charity what many regard as a matter of corrective responsibility.

The economics of responsibility-sharing also shape which legal solutions are viable. A binding obligation to admit displaced persons imposes concentrated, visible costs on receiving states while the benefits---averted humanitarian catastrophe, regional stability---are diffuse and shared. This is a classic collective-action problem, and it predicts precisely the outcome observed: states prefer non-binding commitments and discretionary, capped pathways to enforceable duties of admission. Financial mechanisms that compensate or assist front-line and receiving states---adaptation funding, loss-and-damage arrangements negotiated in the climate regime, and development assistance directed at displacement-affected regions---are therefore not separate from the legal question but integral to it. Protection regimes that ignore the distribution of costs will founder on the unwillingness of states to absorb them.

Geopolitically, large-scale displacement reshapes the strategic environment in ways that give even reluctant states an interest in orderly management. Disorderly mass movement strains borders, fuels domestic political reaction, and can destabilize regions already under pressure. The Falepili Union illustrates how a larger power may convert humanitarian assistance into strategic advantage, securing influence over a smaller state's external arrangements in exchange for a mobility pathway. This dynamic cuts in two directions. It can mobilize resources and political attention that purely humanitarian appeals cannot, harnessing self-interest to protective ends. But it also threatens to make protection contingent on strategic value, leaving displaced populations of no geopolitical interest to any powerful patron without recourse. A regime that relies heavily on bilateral bargains thus risks reproducing, in the domain of climate displacement, the same hierarchy of strategic worth that governs international politics generally. Recognizing this is essential to designing safeguards---common standards, multilateral funding, and minimum protections---that prevent the most vulnerable from falling between the strategic priorities of the powerful.

CONCLUSION

Ioane Teitiota lost his case. He was returned to Kiribati, and the legal system that considered his claim concluded, correctly on its own terms, that he was not a refugee. The unsatisfying quality of that outcome---legally sound, yet morally jarring---captures the predicament this article has examined. The 1951 Convention is not failing climate-displaced persons through neglect or bad faith; it is doing exactly what it was built to do, which is to protect a specific kind of person fleeing a specific kind of harm. Environmental displacement is simply not that harm, and no amount of interpretive ingenuity will make it so.

The lesson is not despair but realism about where protection will come from. It will not come from a single new treaty, at least not in the foreseeable political climate, and it will not come from forcing climate displacement into a refugee definition that cannot hold it. It will come, if it comes, from the patient development of a layered regime: a non-refoulement floor grounded in the right to life, of which *Teitiota* was the first tentative articulation; regional and bilateral frameworks that reach the internally displaced and create orderly pathways; and anticipatory arrangements for the states that face the loss of territory itself.

Yet the deepest point is that the gap is not, at bottom, a problem of legal categorization. The categories matter, and clarifying them has been the work of this article. But a person displaced by a rising sea needs a place to go and a state willing to receive them, and no definition can conjure either. The law can establish floors, channel movement, and distribute obligation. What it cannot do is generate the political will to share responsibility for a harm that the most vulnerable populations have done the least to cause. That will is the scarce resource. The function of a well-designed legal regime is to make its exercise easier, more orderly, and more just---not to substitute for it.

REFERENCES

African Union. (2009). *African Union Convention for the Protection and Assistance of Internally Displaced Persons in Africa (Kampala Convention).*

Government of Australia & Government of Tuvalu. (2023). *Australia--Tuvalu Falepili Union treaty*.

Human Rights Committee. (2020). *Views adopted by the Committee under article 5(4) of the Optional Protocol, concerning communication No. 2728/2016 (Teitiota v. New Zealand)* (CCPR/C/127/D/2728/2016). United Nations.

McAdam, J. (2012). *Climate change, forced migration, and international law*. Oxford University Press.

Organisation of African Unity. (1969). *Convention governing the specific aspects of refugee problems in Africa*.

United Nations. (1951). *Convention relating to the status of refugees* (United Nations Treaty Series, Vol. 189, p. 137).

United Nations. (1966). *International Covenant on Civil and Political Rights* (United Nations Treaty Series, Vol. 999, p. 171).

United Nations. (1967). *Protocol relating to the status of refugees* (United Nations Treaty Series, Vol. 606, p. 267).

United Nations. (2018). *Global Compact for Safe, Orderly and Regular Migration* (A/RES/73/195).

United Nations High Commissioner for Refugees. (2020). *Legal considerations regarding claims for international protection made in the context of the adverse effects of climate change and disasters*.

PAPER 03 / 08

Ownership of Legacy

Michael Jackson, Sony Music, and the Rights Battle

Armin Vats

ABSTRACT

This paper explores the legal and cultural disputes surrounding Michael Jackson's music catalog, focusing on Sony Music's acquisition of a major stake and the challenges raised by Jackson's family. It examines how art becomes property, how that property becomes contested, and what the conflict reveals about pop culture in the corporate age. By analyzing court rulings, estate structures, and economic stakes, the paper highlights the tension between family legacy and corporate control. It argues that the dispute is not only about money but about how society defines ownership of cultural memory.

KEYWORDS *Michael Jackson, Sony Music, music catalog, estate law, posthumous ownership, cultural legacy*

Michael Jackson was more than an entertainer; he was a cultural force whose music shaped generations and whose influence stretched across continents. His catalog is not merely a collection of songs but a living archive of pop culture, a billion-dollar asset wrapped in melody and memory. After his death in 2009, the question of who controls that archive became one of the most contested issues in entertainment law.

When Sony Music struck a deal to buy half of Jackson's catalog for roughly six hundred million dollars, the headlines focused on the numbers. Behind those numbers, however, lay a deeper conflict between family legacy and corporate control. Katherine Jackson, Michael's mother, argued that the sale violated his wishes, but courts disagreed, ruling that the executors had the authority to proceed. The estate, already entangled in a seven-hundred-million-dollar tax dispute with the Internal Revenue Service, viewed the Sony deal as a means to stabilize and maximize value.

This topic merits study as more than a legal battle, because it reflects how society treats art once the artist is gone. The central question is who gets to decide what happens to the music: the family that lived with the artist, or the corporation that distributes him to the world. That tension is what makes the case worth examining. This paper first outlines the background of Jackson's estate and Sony's long history with his music. It then examines the legal disputes, the economic stakes, and the perspectives of both the family and Sony, before discussing the broader implications for pop culture, ownership, and legacy.

Understanding the dispute requires examining how Jackson's estate was structured and how Sony became tied to his music. Jackson's will directed his assets into the Michael Jackson Family Trust, with his children as the main beneficiaries and Katherine Jackson as a life beneficiary of a subtrust. The executors were granted broad authority to manage and sell assets. Sony, for its part, had distributed Jackson's music since the 1980s through record deals, licensing, and global distribution, and its interest in acquiring the catalog was both financial and strategic.

After Jackson's death, the estate became one of the most profitable in entertainment history, with annual revenues reaching hundreds of millions of dollars driven by streaming, licensing, and posthumous releases. Sony recognized an opportunity to secure long-term control of the catalog and thereby strengthen its position in the global music market.

THE LEGAL DISPUTES AND COURT RULINGS

The legal battles centered on whether the executors had the right to sell Jackson's catalog and whether the sale violated his wishes. Katherine Jackson argued that Michael would not have wanted his catalog sold, and lawsuits were filed claiming the sale ran contrary to his will. Courts rejected these claims, ruling that the executors held broad discretion under the will and consistently upholding their decisions.

Complicating matters, the estate was locked in a seven-hundred-million-dollar battle with the Internal Revenue Service that affected how funds flowed into the family trust. The rulings ultimately showed that legal authority rested with the executors rather than the family, creating tension between emotional claims of legacy and legal claims of property rights.

ECONOMIC STAKES AND CATALOG VALUATION

The Sony deal concerned not only ownership but also control of distribution, licensing, and future profits. Jackson's catalog includes songs such as "Thriller," "Billie Jean," and "Beat It," which continue to generate millions of dollars annually. The deal valued the catalog at more than one billion dollars, with Sony paying roughly six hundred million for a fifty percent stake, and annual revenues from Jackson's music remain among the highest for any deceased artist.

For Sony, the purchase was strategic, securing dominance in the global music market and providing leverage in streaming, film licensing, and advertising. For the family, by contrast, the deal felt like losing control of Michael's legacy to a corporation.

PERSPECTIVES: FAMILY VERSUS SONY

From the family's perspective, the catalog is not simply property but memory, legacy, and Michael's voice; selling it feels like selling him. Katherine Jackson argued that Michael would not have wanted his music controlled by a corporation, and the family regards the estate as a guardian of his wishes rather than a business venture.

From Sony's perspective, the catalog is a global asset that the company is best positioned to preserve, distribute, and expand. Ownership ensures that the music continues to reach new markets, films, and streaming platforms, and the company frames the deal as a matter of sustainability and expansion rather than mere profit.

DISCUSSION: POP CULTURE, OWNERSHIP, AND LEGACY

This dispute illustrates how art becomes contested property once the artist is gone. On one side, the family seeks to protect memory and legacy; on the other, the corporation seeks to expand reach and profits. Both positions carry weight: the family speaks from emotion and loyalty, while the corporation speaks from strategy and sustainability. The case also raises the question of whether artistic heritage should be treated primarily as a financial asset, because Sony's deal ensures global distribution yet highlights how money determines what survives.

The broader implications for pop culture are significant. Fans often feel disconnected when corporations control an artist's legacy, family disputes can create uncertainty about how a legacy is managed, and corporate ownership can bring stability while risking the commodification of art. Pop culture itself becomes a battleground between memory and market.

SONY'S HISTORY OF ARTIST DISPUTES: FROM GEORGE MICHAEL TO EMINEM

Jackson's estate battle was not the first time Sony faced accusations of prioritizing corporate control over artistic freedom. Examining other disputes places the case in context and shows that Sony's approach has often been perceived as heavy-handed.

GEORGE MICHAEL V. SONY (1992--1994)

George Michael, one of the biggest pop stars of the early 1990s, sued Sony in the United Kingdom, arguing that his contract amounted to an unreasonable restraint of trade. He claimed that Sony had deliberately under-promoted his album Listen Without Prejudice Vol. 1 because he refused to market himself as a sex symbol, and he described the contract as professional slavery, highlighting how multi-album deals locked artists into long-term obligations while granting labels disproportionate control. The High Court ruled against him, but the case became a landmark in discussions of artist rights and corporate power (Panayiotou v Sony Music Entertainment, n.d.).

EMINEM AND DIGITAL ROYALTIES (2000S--2010S)

Eminem's publishers sued over how Sony and its affiliates classified digital downloads, with the dispute centering on whether downloads should be treated as sales, which carry a lower royalty rate, or as licenses, which carry a higher one. Eminem's representatives argued that digital distribution differed fundamentally from physical sales and that artists deserved a larger share. The case set an important precedent for how musicians are paid in the digital era, compelling labels such as Sony to adjust their contracts.

OTHER SONY ARTIST CLASHES

Other artists experienced comparable friction. Kelly Clarkson publicly criticized the Sony executive Clive Davis for attempting to control her artistic direction, particularly around her album My December, and Shakira disputed contract terms in the early 2000s before moving to other labels for greater creative freedom. Although less dramatic than Jackson's dispute, these cases reinforce the perception that Sony frequently pushes artists toward commercial strategies that conflict with their creative vision.

WHY THIS MATTERS FOR THE JACKSON CASE

Viewed alongside George Michael, Eminem, and others, Jackson's estate dispute appears not as an isolated incident but as part of a pattern in which artists or their representatives accused Sony of prioritizing profit over creativity, control, or legacy. This pattern strengthens the argument that the Jackson family's objections were not merely emotional but part of a broader critique of how corporations handle art. Sony's defense---that it ensures global distribution and sustainability---is valid, yet history shows that artists often feel sidelined in the process.

CULTURAL IMPACT OF JACKSON'S MUSIC

Jackson's catalog is more than a financial portfolio; it is a cultural archive that shaped generations and influenced dance, fashion, and political expression. "Thriller" revolutionized the music-video industry by turning videos into cinematic events, and "Billie Jean" introduced iconic dance moves that became global symbols of pop culture. His music frequently carried social messages, from racial unity in "Black or White" to environmental awareness in "Earth Song," while fashion choices such as the single glove, military jackets, and fedora became cultural trademarks. His catalog continues to appear in films, commercials, and sporting events, demonstrating that his influence is not frozen in the past but constantly reemerging in new contexts.

COMPARATIVE CASE STUDIES: OTHER ARTISTS' ESTATES

Jackson's estate dispute is not unique, as other artists have faced similar battles over control of their legacies. Prince died without a will, leaving his estate in disarray and prompting family disputes that delayed catalog releases and created uncertainty for fans. Elvis Presley's estate balanced family wishes with corporate partnerships, and Graceland became a commercial tourist site that raised questions about the commodification of memory. In the case of Whitney Houston, posthumous hologram tours sparked debate over whether technology honored or exploited her legacy. These comparisons show that Jackson's case belongs to a larger pattern in which families often fight to preserve memory while corporations push for expansion and profit.

PHILOSOPHICAL ANGLE: OWNERSHIP OF MEMORY

The Jackson dispute raises a deeper question of whether cultural memory can be owned, since music is not only property but shared heritage. Treating songs as assets risks reducing art to numbers; families argue that legacy should be protected rather than monetized, while corporations contend that ownership ensures preservation and distribution. This tension reflects a broader philosophical debate over whether art is a commodity or a form of collective memory.

FAN PERSPECTIVE AND GLOBAL LEGACY

Fans play a crucial role in shaping how a legacy is remembered, and many felt disconnected when Sony gained control, arguing that corporations cannot capture the intimacy of fandom. Their reactions included petitions and online campaigns against corporate ownership, social-media debates about whether Jackson's wishes were respected, and fan-led preservation projects such as archives and tribute concerts. Globally, Jackson's music resonates across Latin America, Africa, and Asia; Sony's control ensures distribution in these markets, yet fans often worry that corporate decisions prioritize profit over authenticity.

ETHICAL QUESTIONS IN POSTHUMOUS OWNERSHIP

The dispute also raises ethical dilemmas about whether corporations should profit from art after an artist's death. Families may feel betrayed when corporate deals override personal wishes, corporations argue that they provide stability and sustainability, and fans question whether art should be commodified once the artist is gone. This ethical debate is central to understanding why Jackson's case matters beyond economics.

CORPORATE POWER AND ARTISTIC FREEDOM

The Jackson estate dispute forms part of a larger conversation about how corporations shape artistic freedom, because Sony's history shows that contracts and catalog ownership often tilt power toward the company rather than the artist. Long-term contracts lock artists into obligations that can last decades and limit their ability to experiment, marketing strategies often prioritize commercial appeal over creative risk, and disputes over royalties, ownership, and distribution reveal how corporations define the value of art in financial rather than cultural terms. For Jackson, this tension was magnified after his death, as the estate became a battlefield where corporate interests sought stability and expansion while family members argued for memory and authenticity. The broader issue is whether corporations should dictate how art is remembered, or whether cultural heritage deserves protection beyond market logic.

LEGACY IN THE DIGITAL AGE

The digital era adds new layers to the debate, because streaming platforms, social media, and virtual performances mean that an artist's work never truly fades but is constantly repackaged and reintroduced to new audiences. Streaming ensures global access while centralizing power in corporate hands, social media allows fans to shape narratives and build grassroots movements that challenge corporate control, and virtual concerts and holograms blur the line between tribute and exploitation. Jackson's legacy illustrates how digital tools can both preserve and distort memory: while Sony ensures distribution, fans often feel that corporate decisions strip away intimacy and turn cultural heritage into a commodity.

CONCLUSION

The dispute between Sony and the Jackson estate underscores the need for clearer frameworks in managing posthumous intellectual property, highlighting the balance between honoring an artist's legacy and maximizing economic value. For pop culture, it is a reminder that behind every song we stream lies a battle over ownership, memory, and meaning. Jackson's case is not unique, but it is among the most visible, and it shows that the death of an artist does not end their influence so much as open new battles over who controls it. Whether resolved in favor of family or corporation, the decisions made shape not only the artist's legacy but also the way society remembers and consumes culture.

REFERENCES

Court approves Michael Jackson catalog sale to Sony despite Katherine Jackson's objections. (n.d.). *Billboard*. https://www.billboard.com/pro/michael-jackson-catalog-sale-sony-court-approves/

Leaving Neverland: Trouble for Michael Jackson's estate? (2019). *Variety*. https://variety.com/2019/music/news/leaving-neverland-michael-jackson-estate-250-million-sony-music-trouble-1203156657/

Michael Jackson estate sells major catalog stake to Sony Music. (2024). *Variety*. https://variety.com/2024/music/news/michael-jackson-estate-sony-music-catalog-sale-1235912345/

Michael Jackson estate sells music rights to Sony at \$1.2B valuation. (n.d.). *Billboard*. https://www.billboard.com/business/business-news/michael-jackson-estate-sells-music-rights-sony-valuation-1235604155/

Michael Jackson estate wins appeal over supposedly fake vocals on posthumous release. (n.d.). *Billboard*. https://www.billboard.com/music/music-news/michael-jackson-estate-wins-appeal-supposedly-fake-vocals-posthumous-release-8472706/

Michael Jackson's estate signs sweeping contract with Sony. (2010, March 16). *The New York Times* (DealBook). https://dealbook.nytimes.com/2010/03/16/michael-jacksons-estate-signs-sweeping-contract/

Panayiotou v Sony Music Entertainment (UK) Ltd. (n.d.). In *Wikipedia*. Retrieved June 26, 2026, from https://en.wikipedia.org/wiki/Panayiotou_v_Sony_Music_Entertainment_(UK)_Ltd

PAPER 04 / 08

The Economics of Dignity

Repugnant Markets and the Moral Limits of Exchange

Avni Verma

ABSTRACT

Standard economics begins with a confident rule: if two adults agree to a trade and both walk away better off, the trade is good and should be allowed. Yet every society forbids a stubborn set of exchanges that fit this description perfectly, from selling a kidney to selling a vote. This paper asks why, and argues that the answer is best captured by an old word that economics long ignored, dignity. Drawing on Alvin Roth's idea of repugnant markets, Michael Sandel's distinction between the fairness and corruption objections to commerce, and the experimental literature on motivation crowding, the paper makes a case that dignity is not an irrational intrusion into the market but useful information about where the market's ordinary logic stops working. It shows that introducing a price can change the very thing being sold, that the boundary of what feels repugnant shifts over time without becoming meaningless, and that in genuinely hard cases such as the shortage of transplant organs, dignity and efficiency truly collide. The paper concludes that a mature economics neither dismisses these limits as sentiment nor treats every market as forbidden, but reads dignity as a map of the market's edges, marking the goods that should be handled with care rather than a price tag.

KEYWORDS *repugnant markets, commodification, motivation crowding, moral limits of markets, dignity, organ markets*

THE ECONOMICS OF DIGNITY:

Here is a strange fact about almost every country on Earth. You are allowed to give one of your kidneys to a stranger, a generous and entirely legal act. You are not allowed to sell that same kidney for money. The surgery is identical. The kidney is identical. The only thing that changes between the lawful version and the criminal one is whether cash changed hands.

Somehow, adding a price to a thing the law otherwise permits turns it into something the law forbids.

That should puzzle an economist. The discipline's starting assumption is that voluntary trade is good. If two adults agree to an exchange and both expect to come out ahead, then stopping them seems to destroy value for no reason. By that logic a person who wants to sell a kidney and a patient desperate to buy one should be left to make their deal, both better off than before. Yet we ban it, and we ban a whole family of trades like it, and most people feel the ban is right even if they cannot quite say why.

The philosopher Immanuel Kant drew a line that names what is going on. Some things, he wrote, have a price, which means they can be swapped for an equivalent. Other things have what he called dignity, which means they admit of no equivalent at all (Kant, 1785/1998). The claim of this paper is that economics, for most of its history, tried to put everything in the first category, and that the things which refuse to go there are not an embarrassment to be explained away. They are information. The discomfort people feel when certain goods are priced is a signal, and it marks the places where the ordinary logic of the market quietly stops working.

THE PUZZLE OF THE FORBIDDEN TRADE

To feel the puzzle properly, start with the logic the ban seems to violate. When two people trade freely, each gives up something they value less for something they value more, so both are better off the instant the deal is done. No one outside the trade is harmed. On this reasoning, blocking a willing buyer and a willing seller is simply throwing away gains that hurt

nobody. It is one of the cleanest and most powerful arguments in all of economics, and it underlies the case for free markets everywhere.

And yet the forbidden trades are everywhere too. You cannot legally sell your vote, though it is yours. You cannot sell yourself into slavery, even willingly. In most places you cannot sell sex, or a baby, or a seat on a jury, or a passport. The economist Alvin Roth, who helped design the real system that matches kidney donors to patients, gave these their name. He called them repugnant markets, transactions that some people want to carry out and that other people do not want them to, not because of any harm to outsiders but because the exchange itself strikes observers as wrong (Roth, 2007). The defining feature is exactly what makes them puzzling. The objection is not that a third party gets hurt. The objection is to the trade as such.

So the question sharpens. If the buyer wants it, and the seller wants it, and no bystander is injured, what is left to object to? Answering that honestly is the whole task, and it turns out there are two very different answers hiding inside the single feeling of repugnance.

TWO REASONS WE BALK

The political philosopher Michael Sandel pulled these two answers apart, and the distinction is the single most useful tool in this whole debate (Sandel, 2012). When we recoil from a market, we are usually making one of two complaints, and they are not the same.

The first is the fairness complaint. A trade is only truly voluntary if the person agreeing to it is genuinely free to refuse, and desperation is not freedom. When a man living in poverty sells his kidney to feed his family, the contract may be signed willingly in a narrow sense, but the willingness is manufactured by need. The buyer is not really bargaining with a free equal. He is taking advantage of how little choice the seller has. On this view the problem is not the kidney sale in the abstract but the gross inequality that turns consent into something closer to coercion.

Fix the desperation and the objection would soften.

The second complaint is different and deeper. Sandel calls it corruption, and it says that some goods are degraded, changed for the worse, by the very act of being priced. A friendship you pay for is not a friendship. An apology delivered by a hired stand-in is not an apology. A

prize you bought is not the honor that the same prize would be if it were won. Here the trouble is not unfairness in the bargain. You could imagine a perfectly fair, fully informed sale and the objection would remain, because putting the thing up for sale empties it of the meaning that made it valuable in the first place. The market does not just allocate the good. It alters it.

Keeping these two apart matters, because they point to different remedies and different limits. The fairness objection is about the conditions around a sale and can in principle be answered by reducing inequality or regulating the terms. The corruption objection is about the nature of the good itself, and no amount of fair bargaining can cure it. Much confused argument about markets comes from mixing the two together. Dignity, it turns out, lives mostly in the second.

WHEN A PRICE CHANGES THE THING

The corruption idea can sound like soft philosophy until you notice that economists have measured it. There is now a substantial body of evidence that introducing money into a setting can crowd out the motives that made the setting work, sometimes leaving everyone worse off.

The phenomenon even has a name, motivation crowding, and a few famous demonstrations make it vivid.

The classic case is blood. In a celebrated study of donation systems, the social scientist Richard Titmuss argued that paying people for blood could actually reduce both the quantity and the quality of supply, because payment stripped the act of its character as a gift and drove away the altruistic donors who gave precisely because it was unpaid (Titmuss, 1970). Turn a noble act into a small job and you do not simply add a cash incentive on top of the good feeling. You can replace the good feeling, and lose the donors who came for it.

An even sharper example comes from a daycare in Israel, studied by the economists Uri Gneezy and Aldo Rustichini in a paper bluntly titled "A Fine Is a Price" (Gneezy & Rustichini, 2000). The daycare was troubled by parents arriving late to collect their children, so it introduced a fine for lateness, expecting the penalty to discourage it. Lateness went up. By attaching a price to being late, the daycare had quietly converted a moral failing into a service you could simply

buy. Parents who had felt guilty about inconveniencing the staff now felt they were paying for extra time, fair and square, and so they did it more. Worse, when the fine was later removed, the lateness did not return to its old level. The sense of moral obligation, once priced, did not come back for free.

The lesson running through these cases is that money is not the neutral, frictionless tool the basic model assumes. A price carries its own meaning, and that meaning can displace the social and moral motives already at work. This is not a mystical claim. It is a measurable effect with real consequences for supply, and it gives the corruption objection something the fairness objection always had, namely hard evidence. Sometimes pricing a good really does damage it, and the damage shows up in the numbers.

THE EVERYDAY VERSIONS

It would be easy to think all this lives only in dramatic territory like organs and votes, but the same logic runs through small, ordinary choices. Consider paying to skip a queue. Many of us feel a flicker of unease at the idea of a rich person buying their way to the front of a line at a clinic or a public office, even though the transaction is voluntary and the line moves at the same speed for everyone behind. The discomfort is the dignity intuition in miniature. A queue rations by patience, treating everyone's time as equal, and converting it into a thing money can jump quietly changes what the line means, from a fair wait into one more thing the wealthy can opt out of.

Or consider schemes that pay children to read books or to get good grades. They sometimes work in the short run, but they carry the familiar risk: a child paid to read may learn that reading is a chore worth doing only for cash, and lose the intrinsic pull that the payment was meant to encourage. The same shadow falls over paying people to exercise or to vote. In every case the worry is not that the incentive fails to motivate but that it can teach the wrong lesson about why the thing is worth doing, crowding out a better motive with a thinner one. These low-stakes examples matter because they show that dignity is not a rare exotic constraint reserved for life-and-death markets. It is a feature of how meaning and money interact all the time, in places we barely notice.

A skeptic has a strong reply ready at this point. Is repugnance not just taboo, a bundle of squeamish feelings that societies eventually outgrow? History seems to support the worry.

Charging interest on a loan was once condemned as the sin of usury, morally unthinkable, and is now the ordinary business of every bank. Life insurance was once attacked as a repugnant bet on human death, a way of putting a crude price on a life, before it became a respectable and even responsible purchase (Zelizer, 1979). The line that marks the forbidden clearly moves. Why treat a feeling that shifts with the century as if it carried any real information?

The answer is that a boundary can move without being arbitrary. A coastline shifts too, yet it still reliably tells you where the land stops and the sea begins. The fact that societies have relocated the edge of the market over time does not show the edge is imaginary. It shows that the edge tracks conditions that themselves change, such as how well we can regulate a trade, how desperate the likely sellers are, and whether a given good can survive being priced. When life insurance was poorly understood and lightly regulated, treating it as a grim wager was not foolish. As institutions matured and the practice was reframed as protecting a family rather than betting on a death, the repugnance faded for good reasons, not merely because people got used to it.

That is the heart of the argument. Dignity is not a superstitious barrier that a clever economist should learn to ignore. It is a rough, inherited signal pointing at the cases where the market's normal promise, that voluntary trade makes everyone better off, is most likely to fail, either because consent is hollow or because pricing destroys the good. Read that way, repugnance is doing a job. It is flagging the transactions that deserve a second look rather than an automatic green light. The mistake is not to feel it. The mistake is to treat the feeling as the end of the analysis instead of the start of one.

THE HARD CASE: KIDNEYS

None of this dissolves the genuinely difficult cases, and honesty requires facing the hardest one directly. Thousands of people die every year waiting for a kidney that never comes, because the supply of donated organs falls far short of the need. A regulated, legal market in kidneys might close that gap and save many of those lives. The efficiency argument here is not a cartoon to be swatted away. It is real, it is humane, and the lives it points to are real lives. Iran, which permits a regulated system of paid kidney donation, is often cited as the one country without a deadly waiting list for the organ.

Against that stands everything this paper has described. A kidney market would draw most of its sellers from the poor, turning a grim choice born of desperation into a standing feature of the economy, and it risks treating the body itself as inventory. The fairness objection and the corruption objection both bite hard. So we are left with a true collision, lives that might be saved on one side and human dignity on the other, and no clever reframing makes the tension vanish. Anyone who tells you this case is easy is selling something.

What the dignity lens offers is not a tidy answer but a better way to hold the question. It tells us that if such a market is ever allowed, the things to guard most fiercely are precisely the ones the two objections identify: protect sellers from desperation with floors and safeguards, regulate the terms so the exchange is not raw exploitation, and design the system so that a kidney does not become just another commodity on a shelf. The point of taking dignity seriously is not that it forbids every hard trade. It is that it tells us what must be protected even when we proceed, and what we will have lost if we forget to.

CONCLUSION

Economics earned its power by assuming that almost everything has a price, that goods can be compared, traded, and allocated by willing exchange. For most purposes that assumption is a triumph. But there has always been a remainder it could not absorb, the things people flatly refuse to sell and resent seeing sold, the cases where adding money to a willing trade turns it into a crime or an insult. For a long time the discipline treated that remainder as noise. The argument here has been that it is signal.

Dignity, in this reading, is the name for whatever it is that certain goods possess that makes pricing them either exploit the seller or destroy the thing itself. It is not a wall around the market built out of sentiment, and it is not a license to forbid every exchange that makes someone uncomfortable. It is closer to a map of the market's edges, drawn over centuries of trial and error, marking the places where the ordinary machinery should slow down and proceed with care. The mature response is neither to bulldoze those edges in the name of efficiency nor to treat them as sacred and fixed, but to read them as data about where and why markets fail to deliver what they promise.

Kant said that what has dignity admits of no equivalent. He meant it as a claim about morality, but it doubles as surprisingly good economics. The goods that resist a price are exactly the ones a wise market handles gently, or leaves alone. Knowing which those are, and why, is not a retreat from economic thinking. It is one of its more demanding and humane achievements.

REFERENCES

Gneezy, U., & Rustichini, A. (2000). A fine is a price. *The Journal of Legal Studies, 29*(1), 1--17. Kant, I. (1998). *Groundwork of the metaphysics of morals* (M. Gregor, Trans.). Cambridge

University Press. (Original work published 1785)

Radin, M. J. (1996). *Contested commodities*. Harvard University Press.

Roth, A. E. (2007). Repugnance as a constraint on markets. *Journal of Economic Perspectives, 21*(3), 37--58.

Sandel, M. J. (2012). *What money can't buy: The moral limits of markets*. Farrar, Straus and Giroux.

Titmuss, R. M. (1970). *The gift relationship: From human blood to social policy*. Allen & Unwin.

Zelizer, V. A. (1979). *Morals and markets: The development of life insurance in the United States*. Columbia University Press.

PAPER 05 / 08

Who Pays When Cash Disappears

The Quiet Economics of Going Cashless

Anya Varma · Lotus Valley International School, Noida Extension

ABSTRACT

The move away from physical cash is usually sold as pure convenience, faster checkouts and no more fumbling for change. This paper argues that convenience is the least interesting part of the story, and that going cashless is really a transfer, quietly moving costs and power from some groups to others. It works through who gains and who loses across four fronts. Consumers gain speed but surrender a free public payment option and a layer of privacy. Merchants trade the headache of handling cash for a per-swipe fee paid to card networks and banks. The unbanked, the elderly, and the poor risk losing access to ordinary commerce altogether. And central banks gain a striking new power, the ability to push interest rates below zero, that physical cash currently blocks. Using the contrasting paths of Sweden, which nearly abolished cash and then got nervous, and India, whose 2016 demonetization forced the question overnight, the paper shows that a cashless economy is not simply a more efficient version of the one we have. It is a different distribution of who pays, who profits, and who decides. It closes by arguing that a public digital cash option, a central bank digital currency designed with care, is the way to keep the gains without quietly taxing the people least able to afford it. *Keywords: cashless economy, financial inclusion, interchange fees, central bank digital currency, demonetization, monetary policy*

Who Pays When Cash Disappears: The Quiet Economics of Going Cashless Cash does something strange the moment you stop noticing it. A ten-rupee note or a dollar bill works for everyone, charges no fee, needs no battery, asks no permission, and remembers nothing about you. It is the only form of money a person can use without a bank, a phone, a signal, or anyone else's approval. We are now busily replacing it, tap by tap, and the pitch is always the same single word: convenience. This paper is about everything that word hides.

The claim here is simple to state and easy to overlook. Going cashless is not just a faster version of the economy we already have. It is a quiet transfer. Every time a payment moves from cash to card or app, a small bundle of costs, risks, and powers shifts from one group of people to another, and the groups gaining are rarely the same as the groups paying. Convenience is real, but it is the wrapping, not the package. Underneath it sits a redistribution that almost nobody voted for and that mostly runs in one direction. To see the transfer clearly, it helps to stop asking the usual question, is cashless good or bad, and ask a sharper one instead. Good or bad for whom? The rest of this paper takes that question to four groups in turn: the shopper at the counter, the merchant behind it, the people with no bank account at all, and the central bank sitting above the whole system. Two real episodes, one in Sweden and one in India, then show what happens when a society actually pushes cash toward the exit. The picture that emerges is not an argument against digital money. It is an argument for being honest about its bill, and for deciding on purpose who should pay it.

THE SHOPPER: SPEED NOW, SOMETHING GIVEN UP LATER

Start at the checkout, because that is where the convenience is most obvious and most genuine. Tapping a card or a phone is faster than counting out notes, and faster checkouts add up to real time saved across millions of transactions a day. No honest account of cashless payment can wave this away. For most people, most of the time, the swap feels like a clear upgrade, and it largely is. The cost to the shopper is quieter and shows up later. Two things get handed over at that counter without any sign saying so. The first is a free public option. Cash is money issued by the state that anyone can use at no charge. Once a shop stops taking it, every purchase you make there now runs through a private company that takes a cut somewhere in the chain. You may not see the fee, but you are now inside a system you pay to use, with no free lane left. The second thing handed over is memory. Cash forgets you the instant it leaves your hand. A digital payment does the opposite. It records who paid whom, when, where, and for what, and that record becomes data owned and used by someone else. None of this makes the shopper worse off at the

moment of purchase. It makes them worse off in ways that are spread out and easy to overlook, which is exactly why they get overlooked. None of this is a reason for a shopper to mourn at the till, and the paper is not asking anyone to. The point is about awareness, not refusal. A person who understands that tapping a card is also a small act of paying a toll and filing a report can still decide the speed is worth it, and will often be right. What is troubling is not the choice but the fact that it is never presented as one. The free, private option simply thins out of existence as more shops go cashless, until one day it is gone and no one remembers being asked. A cost you consent to is a price. A cost that disappears your alternatives before you notice is something else, and it is worth naming the difference.

THE MERCHANT: TRADING ONE COST FOR ANOTHER

The shopkeeper's side of the same counter tells a different story, and it is the one most shoppers never think about. Cash is not free for merchants either. It has to be counted, stored, guarded against theft, and physically carried to a bank, and all of that is cost and risk. When customers switch to cards, that burden lifts, which is a real benefit and part of why many merchants welcome the change. But the burden does not vanish. It changes shape, from a cost the merchant managed themselves into a fee the merchant now pays to someone else on every single sale. That fee has a name worth knowing: interchange, the charge that card networks and banks take from each card transaction. In the United States, card networks' own published rate schedules put the average interchange fee at roughly two percent of the transaction value, and total processing fees paid by merchants have run into the tens of billions of dollars a year (Federal Reserve Board, 2024). On a busy day of small purchases, those slices add up to a meaningful chunk of revenue flowing out of the local business and toward the payment industry. Here is the part that quietly reaches everyone, including people paying cash. A merchant facing a fee on most sales does not simply absorb it. They raise prices a little across the board to cover it. So the cost of running the card system gets baked into the sticker price for everybody, which means even the dwindling group

still paying cash ends up subsidizing the convenience of the people who tap. The fee looks like it falls on the merchant. It actually spreads onto all of us. There is a deeper reason this fee is so hard to escape, and it explains why merchants rarely just refuse cards. Payment networks are what economists call two-sided markets. The network is valuable to shoppers because so many shops accept it, and valuable to shops because so many shoppers carry it. Once almost everyone holds a particular card, a merchant who refuses it loses real customers, so they accept it and swallow the fee even while resenting it. That selfreinforcing pull is why a handful of large networks come to sit in the middle of nearly every transaction, collecting a toll from a position no single merchant can afford to walk away from. The convenience of universal acceptance and the inescapability of the fee are the same fact seen from two sides. The shopper enjoys that the card works everywhere. The merchant pays for the reason it does.

THE LEFT OUT: WHEN THE FLOOR FALLS AWAY

Now move to the group the convenience story forgets entirely, the people for whom cash is not a preference but a lifeline. A cashless economy quietly assumes that everyone has the keys to enter it: a bank account, a smartphone, a reliable signal, a steady identity document, and the comfort to use all of them. Large numbers of people are missing at least one of those keys. Globally, 1.4 billion adults remain unbanked, and within any country the people without an account cluster among the poor, the elderly, rural communities, migrants, and those who simply distrust formal finance for reasons of their own (Demirgüç-Kunt et al., 2022). For these people, the disappearance of cash is not an inconvenience. It is a wall. A person without a bank account cannot tap a card they do not have. An elderly customer uneasy with smartphones cannot pay through an app they cannot navigate. When a shop, a bus, or a government office stops accepting cash, it has not merely modernized. It has quietly locked some people out of an ordinary part of life, and it has done so to exactly the people with the least power to complain. The gains from going cashless flow mostly to those already comfortably

inside the financial system, while the costs land hardest on those already outside it. A change marketed as progress can, for the most vulnerable, run progress in reverse. The exclusion also runs deeper than not owning a phone, because entering the digital financial system usually demands things the poorest lack in the first place. A bank account often requires identity documents, a fixed address, and a minimum balance, and each of those is a hurdle for the homeless, the undocumented, migrant workers, and people on the margins of the formal economy. In many places the gap also has a gender dimension, since women are less likely than men to hold accounts or to control the phone a household owns (Demirgüç-Kunt et al., 2022). Cash asks none of this. It demands no paperwork, no permission, and no proof of who you are, which is precisely what makes it the money of people the system has never bothered to count. To remove it is to make formal identity a precondition for buying and selling, and that quietly disenfranchises the very people who already find formal identity hardest to obtain. What looks like a technical upgrade is, for them, a new barrier to entry dressed as modernization.

THE CENTRAL BANK: A NEW LEVER CASH WAS HIDING

The last group is the least obvious and, for an economist, the most interesting. Removing cash hands central banks a power they do not currently have, and understanding it requires one short detour into how interest rates work. A central bank fights a weak economy by cutting interest rates, which makes borrowing cheaper and spending more attractive. But there is a floor. Rates cannot fall far below zero, because of a simple escape hatch: if a bank charged you to keep money in an account, you would pull it out as physical cash and stuff it under the mattress, where it loses nothing. Cash, sitting quietly in wallets, is what stops interest rates from going deeply negative. It is an exit everyone can take, and it is the central mechanism behind what Rogoff (2016) calls the zero lower bound problem. Close that exit and the floor disappears. In a fully cashless economy there is no mattress to run to, because all money lives inside accounts the system can reach. A central bank could then push rates below zero and effectively charge people for holding money, forcing it out into spending during a slump in a way that is impossible today. To a policymaker fighting a deep

recession, that is a powerful new tool. To a saver, it is an unsettling new vulnerability, the prospect of money that shrinks simply for sitting still, with nowhere safe to hide. Whether this power is a blessing or a danger depends entirely on who wields it and how, but the point for this paper is narrower and harder to dispute. The humble bill in your pocket is currently acting as a limit on what the state can do to your savings, and going cashless removes that limit. That is a transfer of power, from the individual to the institution, hidden inside what looks like a payments upgrade. It is also worth noting how invisible this particular transfer is, because no shopper feels it at the counter. Nobody chooses, while tapping a card for coffee, to weaken their own protection against negative interest rates. The power shifts not through any single decision but through the slow disappearance of an option, and an option you no longer have is one you can no longer use as leverage. This is the pattern that recurs throughout the cashless story: the losses rarely arrive as a bill someone hands you. They arrive as the quiet removal of a fallback, a free payment method, a private one, a blackout-proof one, a saver's escape hatch, each a form of insurance that cash provided automatically and for nothing.

ONE SYSTEM, ONE SWITCH

There is a cost that belongs to no single group because it threatens all of them at once, and it has nothing to do with fairness. It is fragility. Cash has a quiet virtue that only becomes visible when everything else fails: it works with nothing behind it. No electricity, no internet, no servers, no signal. A banknote in a blackout still buys bread. A fully digital payment system has the opposite property. It depends on a long chain of power grids, networks, data centers, and working devices, and every link in that chain is a place the whole thing can break. This turns a convenience into a vulnerability the moment something goes wrong. A power cut, a network outage, a failure at a single payment processor, a cyberattack, or in the worst case a wartime strike on infrastructure could freeze an entire economy that has no cash to fall back on. People could find themselves unable to buy food or fuel, not because they lack money, but because the only channel to spend it has gone dark. A society that keeps cash in circulation

always has a manual override, a way to keep trading when the machines stop. A fully cashless one has thrown that override away. This is why even security planners, not just economists, have grown wary of letting cash vanish completely. Resilience is easy to underrate precisely because it is only ever tested on the worst days, and by then it is too late to print more notes.

SWEDEN: THE COUNTRY THAT ALMOST DID IT, THEN HESITATED

Theory is cheap, so consider a place that walked most of the way down this road. Sweden became one of the most cashless societies on Earth, with cash fading to a tiny share of payments and many shops, cafes, and even churches refusing notes entirely (Sveriges Riksbank, 2021). For a while it looked like a clean success story, the convenience argument vindicated in full. Then the country got nervous, and the reasons are exactly the costs this paper has been describing. Officials and the central bank began to worry aloud about the people being left behind, the elderly and others who still relied on cash and were quietly being shut out. They worried about resilience too. A society that runs entirely on digital payments has a single point of failure. A power cut, a network outage, a cyberattack, or a wartime disruption could freeze the entire economy at once, where cash would have kept working. Sweden started encouraging banks to keep cash services available and reminding citizens that physical money still mattered as a backup (Sveriges Riksbank, 2021). The lesson is striking. The country that came closest to abolishing cash looked over the edge and stepped back, not because the convenience was fake, but because the hidden costs turned out to be real enough to concern a wealthy, well-organized state. If Sweden found the bill worth pausing over, poorer and less prepared societies have even more reason to read it carefully.

INDIA: WHEN THE QUESTION ARRIVED OVERNIGHT

Sweden drifted toward cashlessness over years. India had a version of the question forced on it in a single evening. In 2016 the government suddenly withdrew most of the country's banknotes from circulation, a policy known as demonetization, with stated aims that included curbing black money, choking off counterfeit currency, and pushing the economy toward digital payments. How well it met those aims is still debated: economists have disagreed for years over

how much illicit wealth was actually destroyed rather than simply laundered back into the formal system through gold, real estate, or other people's bank accounts. Whatever verdict one reaches on that question, the policy became an unplanned natural experiment in what happens when cash is yanked away from a population fast. The results map closely onto the transfer this paper describes. Digital payment platforms surged as people scrambled for alternatives, and the long-run push toward digital money got a powerful shove. But the immediate pain fell hardest and most visibly on the groups identified earlier. People without bank accounts, daily-wage workers paid in cash, small vendors, and the rural poor were hit first and worst, queuing for hours and losing income while wealthier, banked, smartphone-owning Indians adapted with far less disruption. The episode showed both halves of the truth at once. A determined state really can accelerate the move to digital money, and the costs of doing so are not shared evenly but land on those least able to bear them. India did not disprove the case for digital payments. It illustrated, at full national scale and in fast motion, who pays when cash is taken away without a careful plan to protect them. Set the two countries beside each other and they frame the whole problem from opposite ends. Sweden is the cautionary tale of drift, a rich society sliding toward cashlessness almost by accident and catching itself only at the edge, worried about the excluded and the fragile. India is the cautionary tale of shock, a state forcing the change overnight and watching the costs fall instantly on the poor. One moved too slowly to notice the question until late, the other too fast to soften the answer. Yet both arrived at the same lesson from their different directions. The destination may well be a largely digital economy, but the path there is strewn with people who can be helped or harmed depending entirely on whether anyone bothered to plan for them. Neither speed nor drift protects the vulnerable on its own. Only deliberate design does, which is the case the final section of this paper takes up.

THE STRONGEST CASE FOR KILLING CASH

Honesty requires pausing here to make the other side's argument as well as it can be made, because it is stronger than convenience alone. Cash has a dark side, and it is not small. Precisely

because it is anonymous and leaves no trail, physical cash is the preferred medium of the hidden economy. It lubricates tax evasion, since income paid in cash is income that can vanish from the view of the state, a mechanism well documented in the literature on the shadow economy (Schneider & Williams, 2013). It powers corruption, bribery, and the drug trade, where a paper trail would be fatal. Large-denomination notes in particular serve mainly to move big sums of illicit value quietly, which is why some economists argue they should be abolished outright. On this view, dragging the economy into the digital light is not a convenience at all. It is a public good, a way to recover lost tax revenue, starve criminal networks of their favorite tool, and make the whole system more honest. This argument deserves real respect, and any fair treatment of cashlessness has to grant its force. Reducing the shadow economy is a genuine prize, and digital payments really do make the specific crimes named above, money laundering, tax evasion, and drug trafficking, harder to conceal and easier to trace. But notice what the argument does and does not prove. It establishes that there are real benefits to less cash. It does not establish that those benefits are free, or that they are shared fairly, which is the whole concern of this paper. A policy can do real good at the top, recovering taxes and frustrating criminals, while doing real harm at the bottom, locking out the poor and the unbanked, at the very same time. The existence of a strong case for going cashless does not cancel the question of who pays for it. It just makes the question more urgent, because a benefit that genuine is exactly the kind that tempts a society to stop counting the costs.

THE SURVEILLANCE QUESTION

One cost named only briefly so far deserves its own examination, because it is the one that grows most quietly and is hardest to reverse. Cash is private by nature. It records nothing and reports to no one. Every digital payment, by contrast, is a small act of self-documentation. It tells some company, and potentially the state, what you bought, where, when, and from whom. Multiply that across every purchase of a person's life and you have assembled something cash

never permitted: a complete, searchable map of how someone lives, drawn from where their money goes. In ordinary times this can feel harmless, even useful, since the data powers conveniences people enjoy. The danger is what the map becomes in the wrong hands or under the wrong rules. A complete record of everyone's spending is a tool of extraordinary power. It can be sold, breached, or turned to purposes the payer never agreed to, and in the hands of a government inclined to monitor or punish, it becomes an instrument of control that the anonymity of cash always denied. This is not an argument that digital payment is sinister. It is a reminder that privacy, once a free and automatic feature of using money, becomes something you must now consciously fight to preserve. The retreat of cash quietly converts a default into a luxury, and luxuries tend to end up reserved for those who can afford them.

KEEPING THE GAINS WITHOUT SENDING THE BILL DOWNWARD

Put the four groups and the two countries together and a single conclusion comes into focus. The problem with going cashless is not the digital part. Faster payments, less theft, and easier records are genuine goods worth having. The problem is that the current path captures those goods for the comfortable while quietly mailing the bill to the merchant, the cash-paying customer, and above all the excluded. The convenience is real. The unfairness is in who pays for it. That framing points toward a fix that keeps the gains and changes the distribution. The deepest trouble with a privately run cashless system is that it has no free public option left in it. Cash used to be that option, money anyone could use at no charge and without permission, and letting it vanish hands the entire payment system to private firms that levy a toll. The answer is not to cling to paper forever. It is to build a public form of digital money to play the role cash once did. This is the idea behind a central bank digital currency, a digital equivalent of a banknote issued by the state, usable by anyone without a commercial middleman taking a cut on every transaction (Bech & Garratt, 2017).

Designed carefully, such a public option could thread the needle this paper has been describing. It could preserve a free, universally accepted way to pay, so that the poor and the unbanked are not locked out as cash retreats. It could spare merchants and, through them, all customers from paying a private fee on every sale. It could be built with real privacy protections rather than treating every purchase as data to be harvested. And by keeping a public form of money in everyone's hands, it could even soften the central-bank power problem, leaving people an option that does not exist when only private accounts remain. None of this is automatic. A digital currency built carelessly could deepen surveillance and centralize power instead of checking it, which is precisely why the design matters more than the slogan. The choice a society faces is not really cash versus digital. It is whether the digital money that replaces cash will be a private toll road or a public commons. That choice is still open, and it should be made deliberately rather than drifted into one tap at a time. It is worth being clear about what this proposal is not. It is not a demand that cash be preserved forever out of nostalgia, and it is not a rejection of digital payment, which has brought real and widely shared benefits. The argument is narrower and, I think, harder to dodge. Whatever replaces cash should keep doing the jobs cash quietly did: serving everyone, charging nothing, surviving a blackout, and minding its own business about what you bought. A privately run cashless system, left to itself, abandons every one of those jobs, because no private firm has a reason to perform them for free. Only a public option has that reason, because serving everyone is the job the public sector exists to do. The question is therefore not whether money will go digital. It already is. The question is whether the public will build itself a seat in the system that results, or hand the whole thing to the toll-collectors and hope for the best.

THREE OBJECTIONS

Three replies to this argument deserve a hearing. The first says the market will sort it out on its own. If enough people are being left behind or overcharged, businesses will spot the unmet demand and serve it, and no public option is needed. There is something to this, but it misreads who is being left behind. The excluded are, almost by definition, the least profitable customers:

the poor, the elderly, the rural, the unbanked. A market chases profit, and there is little profit in serving precisely the people who have least to spend. That is why financial exclusion has proved so stubborn even in rich countries with sophisticated banks. The market is very good at serving the comfortable and reliably poor at serving the marginal, which is exactly the gap a public option exists to fill. The second objection runs the opposite way and worries about the cure. A central bank digital currency, it warns, could become the most powerful surveillance tool ever built, a single ledger of every citizen's spending sitting in the hands of the state, far more dangerous than the scattered data of private firms. This is a serious and correct warning, and it is why this paper keeps insisting that design is everything. A public digital currency built without real privacy protection would indeed be a nightmare, trading a private toll road for a government watchtower. The reply is not to dismiss the fear but to treat it as a binding constraint. The case made here is for a public option built with genuine, enforceable privacy, not for handing the central bank a microscope. A tool that can be built well can also be built badly, and the argument is for the first, not a blind faith that only the first will occur. The third objection is simply that this is all premature, that cash is not actually disappearing and the whole worry is overblown. It is true that cash remains widely used in much of the world and shows no sign of vanishing tomorrow. But the direction of travel is not in much doubt, and the groups most at risk are the ones with the least voice to raise an alarm when the change does come. The time to decide how a transition should work is before it has quietly finished, not after, when the infrastructure is built, the habits are set, and the excluded have already been left on the wrong side of a wall nobody remembers choosing to build. Caution about the timing is fair. Using it as a reason not to think the question through is not.

CONCLUSION

The disappearance of cash feels like something happening to us, a tide of convenience we simply ride. This paper has argued that it is better understood as something being decided, transaction by transaction, mostly without debate. Behind the single word convenience sits a

quiet transfer of cost, risk, and power, and the people gaining are seldom the people paying. Shoppers trade a free, private, permissionless option for speed. Merchants trade their own handling costs for a fee that spreads onto everyone. The unbanked and the elderly risk losing the floor beneath ordinary commerce. And central banks gain a lever over savings that the bill in your wallet currently holds shut. Sweden and India, from opposite directions, deliver the same warning. A society can move fast toward a world without cash, but the costs of doing so are real and they fall unevenly, which is why even an enthusiastic Sweden paused and why a sudden Indian shock hurt the poor first. The honest conclusion is not that digital money is a trap to be refused. It is that a cashless economy is a choice with a price tag, and the price is currently being slipped to the people least able to pay it. Keeping cash's best qualities alive in digital form, through a public option built with genuine care, is the way to take the convenience without quietly taxing the vulnerable for it. Cash may be on its way out. What it stood for, money that belongs to everyone and answers to no toll, is worth carrying forward.

REFERENCES

Bech, M., & Garratt, R. (2017). Central bank cryptocurrencies. BIS Quarterly Review, 55– 70. Bank for International Settlements. Demirgüç-Kunt, A., Klapper, L., Singer, D., & Ansar, S. (2022). The Global Findex Database 2021: Financial inclusion, digital payments, and resilience in the age of COVID-19. World Bank. Federal Reserve Board. (2024). Regulation II: Average interchange fee by payment card network. Board of Governors of the Federal Reserve System. Rogoff, K. S. (2016). The curse of cash. Princeton University Press. Schneider, F., & Williams, C. C. (2013). The shadow economy. Institute of Economic Affairs. Sveriges Riksbank. (2021). Payments in Sweden 2021.

PAPER 06 / 08

Income Inequality and the Kuznets Curve

Does Economic Growth Reduce Inequality?

Anany Srivastava

ABSTRACT

In 1955 Simon Kuznets proposed that income inequality follows an inverted-U path as an economy develops, rising during early industrialisation and falling once the economy matures. The hypothesis, known as the Kuznets curve, became one of the most influential ideas in development economics and offered a reassuring implication: that growth, if sustained, eventually narrows the gap between rich and poor. This paper evaluates whether the evidence supports that claim. It first explains how inequality is measured through the Lorenz curve and the Gini coefficient, then sets out the structural logic behind Kuznets's prediction. It distinguishes the early cross-sectional evidence, which appeared to support the curve, from the later longitudinal evidence, which largely did not, and it examines two serious challenges: the East Asian economies that grew rapidly without the predicted surge in inequality, and the resurgence of inequality across developed economies since roughly 1980, documented most prominently by Thomas Piketty. The paper concludes that the Kuznets curve is best understood not as an economic law but as a description of one historical episode, and that the path of inequality is governed far more by institutions and policy than by the level of development alone. It closes by considering Branko Milanovic's reframing of the pattern as recurring "Kuznets waves."

KEYWORDS *Kuznets curve, income inequality, Gini coefficient, economic development, Piketty, structural transformation*

Few questions in economics carry higher stakes than whether economic growth, on its own, eventually makes a society more equal. If it does, then the central task of development is simply to grow, and distribution will take care of itself in time. If it does not, then growth and fairness are separate goals that may require separate tools, and a country can become far richer while leaving its poor behind. The most famous attempt to answer this question produced a single, elegant prediction that has shaped six decades of debate. This paper sets out that prediction, examines the evidence for and against it, and argues that its optimistic promise does not survive contact with the long-run data.

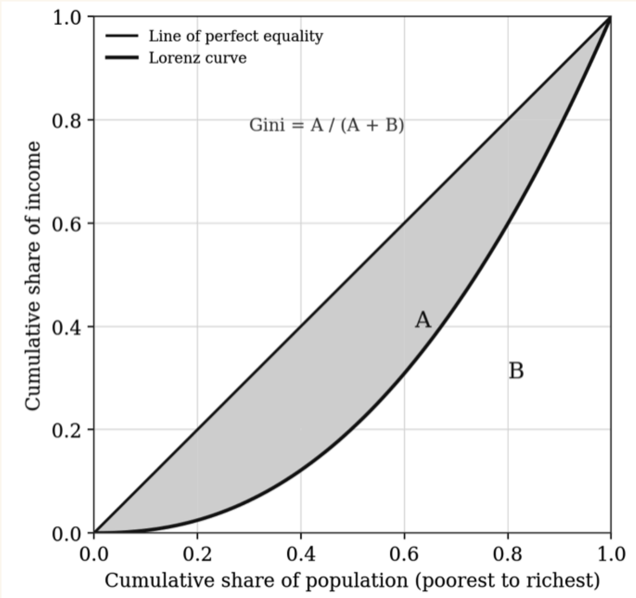
The prediction belongs to Simon Kuznets, who proposed in 1955 that inequality and development are linked by an inverted-U relationship (Kuznets, 1955). In the earliest stages of growth, he suggested, inequality rises, but beyond a certain point continued growth brings it back down, so that a graph of inequality against income per capita traces a hump. The idea is intuitive, visually clean, and politically comforting, and for a generation it was treated almost as a law of development. The question this paper pursues is whether it deserves that status. The answer, drawn from longitudinal data and from the sharp rise of inequality in rich countries since around 1980, is that it does not.

The argument proceeds in five steps. The paper first explains how economists measure inequality, since the whole debate depends on a single number whose meaning must be clear. It then lays out the structural reasoning behind Kuznets's hypothesis. Next it weighs the early evidence that seemed to confirm the curve against the later evidence that undermined it. It then turns to the most serious modern challenge, the return of high inequality in developed economies. Finally it considers how the pattern might be reinterpreted once the simple curve is abandoned.

Before asking how inequality changes with growth, one must be able to put a number on it, and the standard tool is the Gini coefficient. It is built from a simple picture called the Lorenz curve, shown in Figure 2. Imagine lining up a country's households from poorest to richest along the horizontal axis, and plotting on the vertical axis the cumulative share of total income they hold. If income were perfectly equally distributed, the poorest 20 percent would hold 20 percent of income, the poorest half would hold half, and the plot would be a straight diagonal line. Real distributions sag below that diagonal, because the poorer households hold less than their proportional share.

FIGURE 1

The Lorenz Curve and the Gini Coefficient



Note. The diagonal is the line of perfect equality. The Lorenz curve lies below it; the Gini coefficient equals the area between them (A) divided by the total area beneath the diagonal (A + B). A Gini of 0 means perfect equality and 1 means one household holds all income. Figure is illustrative.

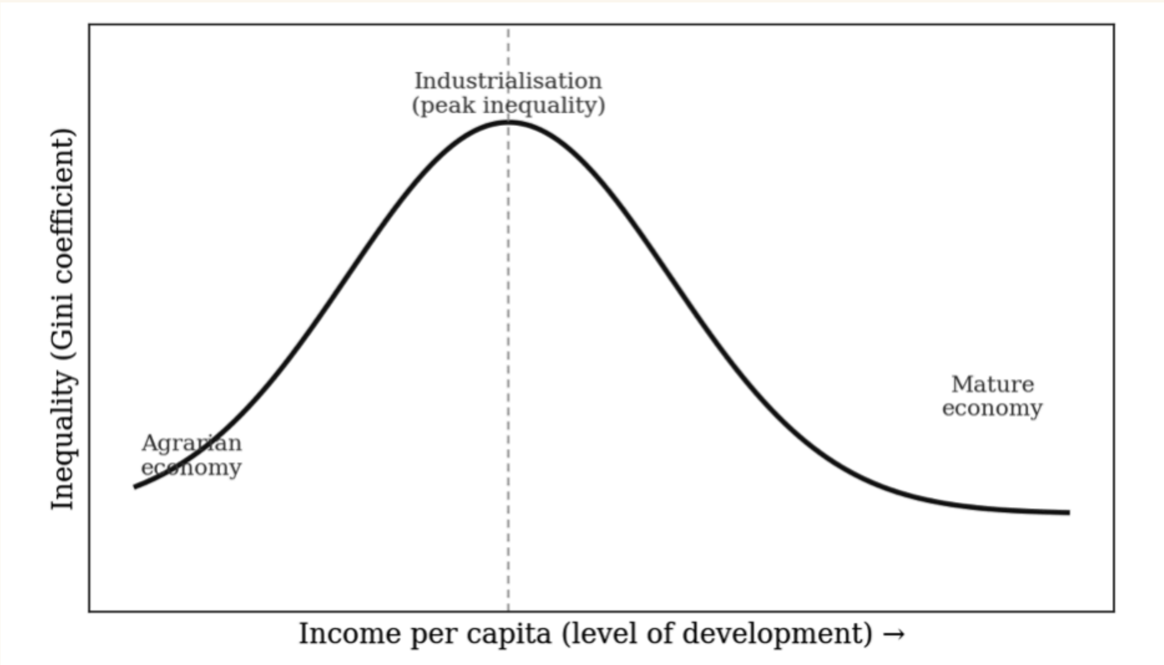
The Gini coefficient captures how far the Lorenz curve sags. It is the area between the line of equality and the Lorenz curve, expressed as a fraction of the whole triangle beneath the diagonal. A Gini of zero describes perfect equality, where everyone has the same income and the curve sits on the diagonal. A Gini of one describes perfect inequality, where a single household holds everything. Real economies fall in between, with most national Gini values lying somewhere between roughly 0.25 and 0.65. The measure is not perfect, since two very different distributions can share a Gini value, but it has the great virtue of compressing an entire distribution into one comparable number, and it is the figure on which the Kuznets debate is conducted.

THE KUZNETS HYPOTHESIS AND ITS LOGIC

Kuznets did not arrive at the inverted-U by accident; he had a structural story for why development should produce it, and the story is worth following because its assumptions are where the hypothesis later proves vulnerable. The shape he predicted is drawn in Figure 1.

FIGURE 2

The Theoretical Kuznets Curve



Note. As an economy develops, the hypothesis predicts that inequality first rises, peaks during industrialisation, and then declines as the economy matures, tracing an inverted U. Figure is a stylised representation of the theoretical relationship.

The mechanism is structural transformation. A poor, pre-industrial economy is mostly agrarian. Incomes are low but also relatively uniform, because most people work the land at similar levels of productivity, so inequality is modest. Industrialisation breaks this uniformity. A modern, high-productivity sector emerges in the cities and pays far more than traditional agriculture, and as some workers move into it while most remain behind, the gap between the two sectors widens the income distribution sharply. Inequality climbs toward the peak of the curve. This is the upswing.

The downswing follows, in Kuznets's account, as the modern sector keeps expanding. Over time more and more of the population moves out of low-paid agriculture and into higher-paid industry, so the share left behind in the poor sector shrinks. Eventually the modern sector dominates, labour grows scarcer relative to capital, wages rise, and the spread of mass education and, later, redistributive taxation and social spending all push inequality back down. A mature economy, on this view, settles at a lower level of inequality than it passed through on the way up. The logic is coherent, and it fit the broad experience of the Western economies Kuznets had in view. The difficulty is that coherence is not proof, and the prediction can be tested.

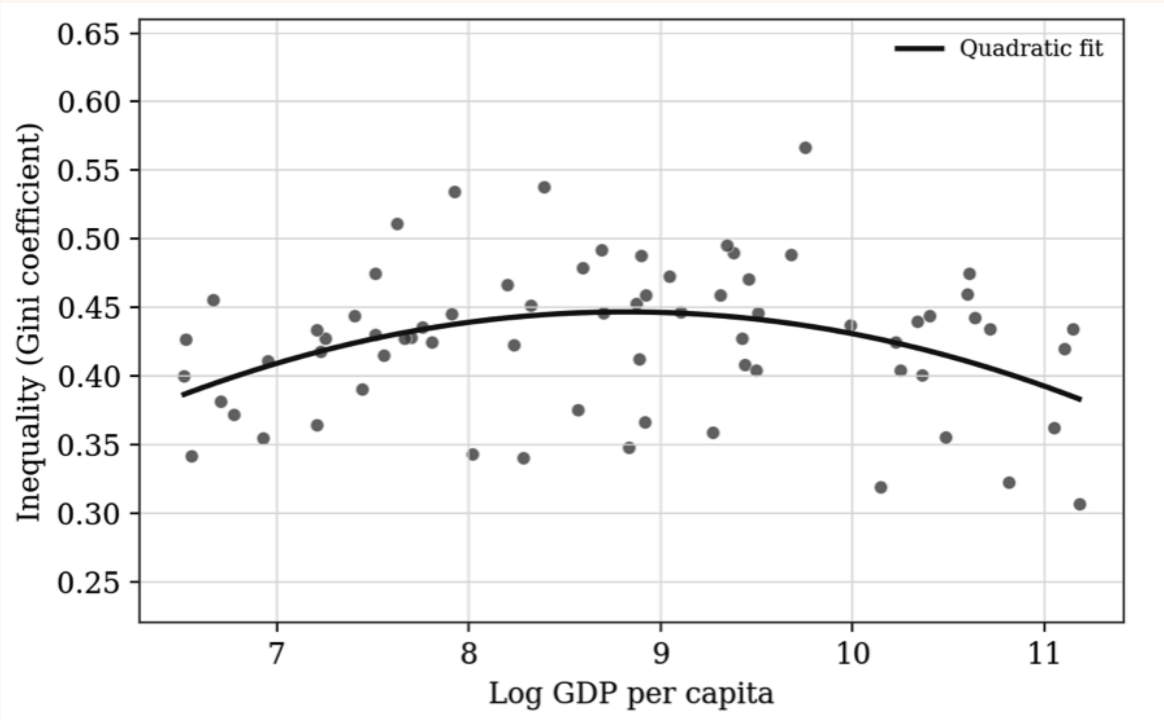
THE EARLY EVIDENCE: A CURVE ACROSS COUNTRIES

The first tests of the hypothesis were cross-sectional. Rather than follow a single country over time, which requires long runs of data that did not then exist, economists compared many countries at a single moment, lining up their inequality against their income per capita. If the Kuznets curve were real, poor and rich countries should both show lower inequality than middle-income countries, producing an inverted-U across the sample. Some early studies did report exactly this pattern, and it lent the hypothesis much of its initial credibility.

Figure 3 illustrates the form such a test takes, plotting Gini coefficients against the logarithm of income per capita across a set of countries. The fitted quadratic curve does bend downward at the ends, hinting at the predicted hump. But the figure also reveals the weakness of this evidence, which is the enormous scatter around the line. Countries at very similar income levels exhibit wildly different inequality, and the curve explains only a small fraction of the variation. The shape can be coaxed out of the data, but it is faint, and it depends heavily on which countries are included.

FIGURE 3

Inequality and Income per Capita Across Countries



Note. Each point represents a hypothetical country. The fitted quadratic bends downward at high incomes, consistent with a weak inverted-U, but the wide scatter shows that development level explains little of the cross-country variation in inequality. Data are illustrative and generated for demonstration.

There is a deeper problem with cross-sectional evidence that no amount of data can fix. Comparing different countries at one moment is not the same as watching one country develop. A poor country today and a rich country today differ in countless ways besides their stage of development, including their history, institutions, and politics, and a pattern across countries need not describe the path any single country actually travels. To test Kuznets's claim properly, one needs to follow the same economies over long stretches of time. When data of that kind finally became available, the curve began to fade.

THE EMPIRICAL CHALLENGE: WHEN THE CURVE FADES

As longer and more consistent inequality series were assembled, researchers could at last test the hypothesis within countries over time rather than across them at a single moment. The results were not kind to the curve. In an influential study assembling a large new dataset on inequality, Deininger and Squire (1996) found little systematic support for the Kuznets relationship once the data were examined country by country. For most economies, the inverted-U simply did not appear in their own historical experience; inequality did not reliably trace the predicted hump as income rose.

The East Asian economies posed an especially sharp challenge. South Korea, Taiwan, and others grew at extraordinary rates through the late twentieth century, precisely the rapid industrialisation that Kuznets's logic says should drive inequality up toward a peak. Yet several of these economies industrialised with comparatively moderate and even stable inequality, achieving what came to be called growth with equity. Land reform, heavy investment in broad-based education, and particular policy choices appeared to let them climb the development ladder without the surge in inequality the curve predicted. If rapid growth can occur without the upswing, then the upswing is not a necessary feature of development, and the mechanism Kuznets described is at most one possibility among several.

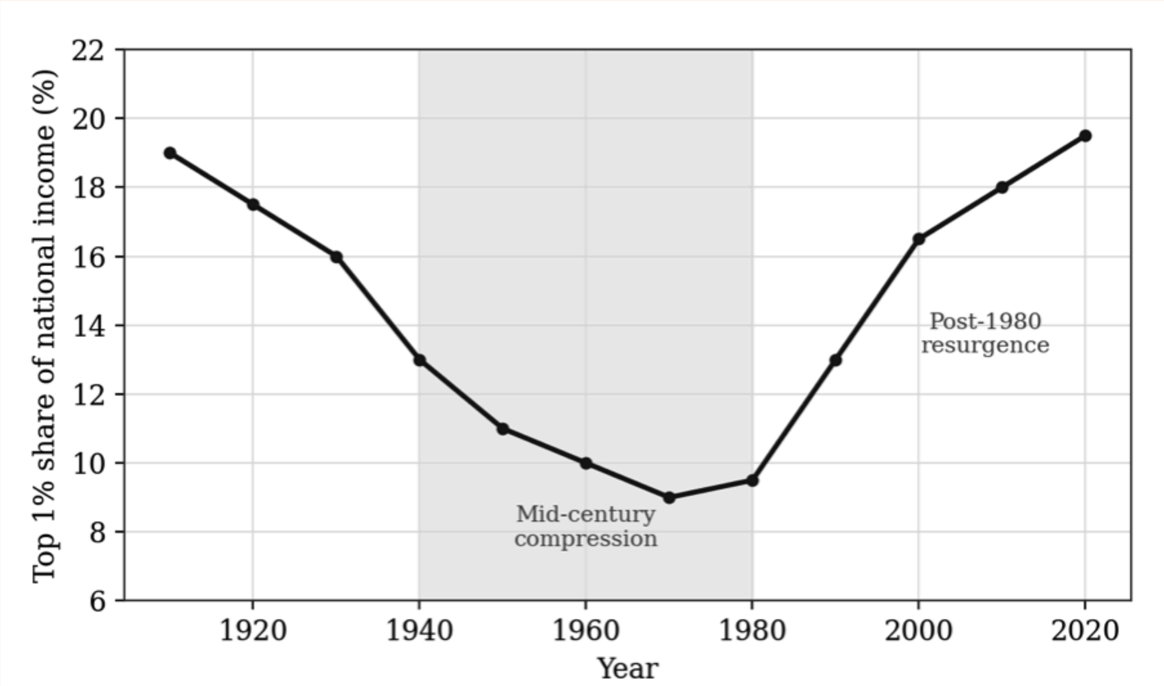
Taken together, the longitudinal evidence and the East Asian cases suggested that whatever regularity the early cross-sectional studies had detected was fragile. The curve was not a reliable description of how inequality evolves as a country grows. It described some histories and not others, which is a very different thing from an economic law. But the most striking blow to the hypothesis came not from the developing world at all. It came from the long-run experience of the rich economies that were supposed to be sitting comfortably on the curve's descending arm.

PIKETTY AND THE RETURN OF INEQUALITY

The Kuznets curve predicts that once an economy matures, inequality settles at a low level and stays there. For roughly the first three decades after the Second World War, the rich economies seemed to confirm this. Inequality in countries such as the United States had fallen to historically low levels and remained compressed, apparently the flat tail of the inverted-U. Then, beginning around 1980, it reversed. Across many developed economies inequality climbed steeply, and by the early twenty-first century the share of income flowing to the very top had returned toward levels not seen since before the curve's supposed descent. Figure 4 sketches this long-run U-shape.

FIGURE 4

Long-Run Top Income Share in a Developed Economy



Note. Stylised long-run path of the top 1 percent's share of national income. Inequality fell to a mid-century low, then rose sharply after about 1980, contradicting the Kuznets prediction that a mature economy remains at low inequality. Series is illustrative and based on the pattern documented by Piketty (2014).

This reversal is the centre of Thomas Piketty's argument in his study of long-run inequality (Piketty, 2014). Drawing on tax records reaching back over a century, Piketty contended that the mid-century compression of inequality was not the natural maturing that Kuznets described, but the product of extraordinary and largely destructive events: two world wars, the Great Depression, and the high taxes and policy choices that followed them, which together destroyed and redistributed concentrations of wealth. Once those forces receded, inequality resumed its climb. On this reading the flat, low tail of the Kuznets curve was an illusion, a temporary trough carved by catastrophe rather than a stable resting place reached by growth.

If Piketty is right, the implication for the Kuznets curve is severe. The hypothesis predicts that development carries inequality down and keeps it there. The actual long-run record of the developed economies is not a hump that levels off but something closer to a U turned the other way, high inequality early in the twentieth century, a mid-century dip, and a renewed rise. Growth did not deliver and maintain equality. Inequality fell when particular shocks and policies pushed it down and rose again when they were withdrawn, which points to causes quite different from the level of development on its own.

REINTERPRETING THE PATTERN

If the simple curve does not hold, what replaces it? One influential answer comes from Branko Milanovic, who proposes that inequality moves not in a single inverted-U but in repeated cycles he calls Kuznets waves (Milanovic, 2016). On this view, each great technological and economic transformation, the first industrial revolution and now the digital and globalised economy, sets off its own upswing and eventual downswing of inequality. The mid-century decline and the post-1980 rise are then not a failed single curve but two phases of successive waves, each driven by forces of technology, globalisation, and politics rather than by a country simply passing a fixed income threshold.

Whether or not the wave metaphor proves durable, it captures the central lesson of the evidence. The path of inequality is not a mechanical function of how rich a country has become. It is shaped by the things the original hypothesis held in the background or left out: the spread of education, the bargaining power of labour, the openness of the economy to trade and capital, the structure of taxation, and the strength of the political coalitions that favour or resist redistribution. Two countries at the same level of development can have very different inequality because they have made very different choices, which is exactly what the wide scatter in Figure 3 was showing all along. Development sets the stage; institutions and policy write the play.

CONCLUSION

Simon Kuznets offered a hypothesis of rare elegance and, in fairness to him, offered it tentatively, aware of how thin his data were. The inverted-U was a reasonable reading of the limited Western evidence available in 1955, and it asked a question that still matters enormously. The trouble is what the idea became in other hands: a near-law of development, carrying the comforting promise that growth would in time dissolve inequality on its own. That promise is what the evidence does not support.

The early cross-sectional studies found only a faint and fragile pattern. The longitudinal data, examined country by country, largely failed to reproduce the curve. The East Asian economies showed that rapid growth need not raise inequality, and the developed economies, by letting inequality surge again after 1980, showed that maturity does not keep it low. The most accurate summary is that the Kuznets curve describes one historical episode reasonably well and the general experience of development poorly. Growth does not reliably reduce inequality. Whether inequality rises or falls depends on what a society does, through its institutions and its policies, far more than on how rich it has become.

This is not a counsel of despair but a clarification of responsibility. If equality were an automatic reward of growth, no one would need to choose it. Because it is not, the distribution of income is revealed as a matter of decision rather than destiny. That is a heavier conclusion than Kuznets's optimistic curve, but it is also a more honest one, and it places the question of who benefits from growth back where it belongs, in the hands of the people and governments who shape the rules.

REFERENCES

Barro, R. J. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth, 5*(1), 5--32.

Deininger, K., & Squire, L. (1996). A new data set measuring income inequality. *The World Bank Economic Review, 10*(3), 565--591.

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review, 45*(1), 1--28.

Lakner, C., & Milanovic, B. (2016). Global income distribution: From the fall of the Berlin Wall to the Great Recession. *The World Bank Economic Review, 30*(2), 203--232.

Milanovic, B. (2016). *Global inequality: A new approach for the age of globalization*. Harvard University Press.

Piketty, T. (2014). *Capital in the twenty-first century* (A. Goldhammer, Trans.). Harvard University Press.

PAPER 07 / 08

Regime Type and Sanctions Resilience

Comparing Democratic and Authoritarian Responses to Economic Coercion

Kabir Srivastava · Sarvottam International School

ABSTRACT

Economic sanctions are among the most frequently used instruments of international statecraft, yet their effectiveness remains highly contested. This article examines how political regime type influences a state's ability to withstand and adapt to economic sanctions. Drawing upon selectorate theory, political accountability, information control, and state capacity, the study argues that authoritarian regimes often possess structural advantages that enhance resilience to sanctions. To evaluate this argument, the article employs a comparative case study approach, analysing Russia following the 2022 invasion of Ukraine, Iran's long-term experience under international sanctions, and South Africa during the anti-apartheid sanctions campaign of the 1980s. The findings suggest that sanctions are generally less effective against authoritarian regimes because political leaders face narrower accountability structures, possess greater control over information environments, and can more effectively shield key supporters from economic costs. However, the South African case demonstrates that sanctions can contribute to meaningful political change when economic pressure reaches influential constituencies, domestic opposition remains active, and international coordination is sustained. The analysis further shows that regime type alone cannot explain sanctions outcomes, as factors such as state capacity, economic adaptation, resource endowments, and alternative international partnerships also shape resilience. The article concludes that the effectiveness of sanctions depends not only on the severity of economic pressure but also on the political and institutional environment through which that pressure is transmitted.

INTRODUCTION

Following Russia\'s 2022 invasion of Ukraine, the country became the subject of the most extensive sanctions campaign in modern history. Policymakers and analysts widely predicted severe economic collapse, widespread political instability, and significant pressure on the Kremlin to change its behaviour. Comparable expectations have accompanied sanctions campaigns against various states throughout the past century. Despite the growing frequency and sophistication of economic sanctions, their effectiveness remains uncertain.

The effectiveness of sanctions remains deeply contested. While some sanctioned governments have modified their policies or entered negotiations, others have exhibited significant resilience, enduring prolonged economic pressure without conceding to the demands of sanctioning states.

Economic sanctions have emerged as one of the most widely used instruments of contemporary international statecraft. Broadly defined as the deliberate restriction of trade, financial, or economic relations with a target state to induce a change in behaviour, sanctions occupy a position between diplomacy and military intervention.

They offer policymakers a means of signalling resolve and imposing costs without resorting to armed force. Consequently, sanctions have become a preferred response to challenges ranging from territorial aggression and nuclear proliferation to democratic backsliding and human rights abuses.

Despite their popularity, the historical record of sanctions presents a mixed picture. Some studies suggest that sanctions rarely achieve their primary political objectives, while others argue that their effectiveness depends heavily on contextual factors such as economic interdependence, multilateral participation, enforcement mechanisms, and the nature of the demands imposed upon the target state. This debate has generated a substantial body of scholarship seeking to explain why sanctions succeed in some circumstances and fail in others.

Among the variables influencing sanctions outcomes, the political character of the targeted state deserves particular attention. Whether a state is governed through democratic institutions or authoritarian structures shapes how economic pressure is experienced, interpreted, and politically managed. Economic hardship does not automatically translate into political change. Instead, sanctions operate through domestic political systems that determine who bears the costs, how information about those costs is communicated, and whether leaders face meaningful accountability for deteriorating economic conditions.

Democratic governments generally operate within institutional environments characterised by electoral competition, political pluralism, independent media, and public accountability. As economic hardship intensifies, citizens have formal mechanisms to express dissatisfaction and pressure political leaders to reconsider costly policies.

Authoritarian governments, by contrast, often possess tools unavailable to democratic systems, including extensive control over information, coercive security institutions, and the capacity to selectively distribute economic burdens across society. These differences suggest that regime type may significantly influence a state's ability to withstand external economic pressure.

This article argues that regime type is among the most important determinants of sanctions resilience, defined here as the capacity of a target state to endure sustained economic sanctions without making the concessions demanded by sanctioning actors. Specifically, it contends that authoritarian regimes frequently possess structural.

advantages that enable them to absorb, redirect, and politically manage sanctions pressure more effectively than democratic systems. However, this advantage

is neither universal nor absolute. Factors such as elite cohesion, access to alternative economic partners, resource endowments, and the international enforcement environment interact with regime type to shape outcomes.

To examine these dynamics, this article adopts a comparative approach centred on three case studies: Russia following the extensive sanctions imposed after the 2022 invasion of Ukraine, Iran's decades-long experience under American and multilateral sanctions, and South Africa during the international sanctions campaign directed against the apartheid regime in the 1980s. Together, these cases illustrate both the strengths and limitations of sanctions as instruments of coercion and provide insight into the political mechanisms that shape resilience and compliance.

By placing regime type at the centre of analysis, this study contributes to ongoing debates regarding sanctions effectiveness, international law, and economic statecraft.

More broadly, it seeks to demonstrate that the success or failure of sanctions cannot be understood solely through measures of economic damage. Rather, it depends on how political institutions transform economic pressure into political consequences, making the study of domestic political structures essential to understanding the real-world impact of economic sanctions.

1. **Selectorate Theory and Political Survival**

The most influential framework for understanding why democracies and autocracies respond differently to economic sanctions is selectorate theory, developed by Bruce Bueno de Mesquita and colleagues (2003). The theory argues that political leaders remain in power by maintaining the support of a coalition of essential supporters, often referred to as the winning coalition. The size and composition of this coalition vary significantly across regime types and fundamentally shape how leaders respond to external economic pressure.

In democratic systems, political leaders depend upon the support of broad electoral constituencies. Because democratic survival is linked to public approval and electoral performance, economic hardship generated by sanctions can quickly become a political liability. Rising inflation, unemployment, declining living standards, and reduced economic growth often translate into voter dissatisfaction, creating pressure on governments to reconsider policies that contribute to international isolation.

Consequently, sanctions can impose both economic and political costs on democratic leaders.

Authoritarian systems operate according to a different political logic. Rather than relying on large electoral majorities, authoritarian rulers typically depend on a relatively narrow coalition consisting of military leaders, party officials, economic elites, or other

influential actors. This structure allows governments to concentrate resources on maintaining elite loyalty while shifting a disproportionate share of economic hardship onto the broader population. As a result, the political consequences of sanctions are often weakened because those most affected by economic decline are not necessarily those whose support determines regime survival.

The implications of the selectorate theory for sanctions effectiveness are significant. Economic sanctions do not directly compel governments to change their behaviour; rather, they seek to generate domestic political pressure that encourages policy concessions. Whether such pressure emerges depends largely on how economic costs are distributed within a political system. In democratic regimes, broad accountability mechanisms increase the likelihood that economic pain will translate into political pressure. In authoritarian systems, narrower support coalitions and greater control over resource allocation frequently allow leaders to insulate themselves from the full political consequences of sanctions.

Selectorate theory, therefore, provides a useful starting point for understanding variations in sanctions resilience. However, it does not fully explain why some authoritarian regimes remain highly resilient while others do not, nor why certain democracies demonstrate unexpected resistance to external pressure. Additional factors, including political accountability, information control, and state capacity, must also be considered.

POLITICAL ACCOUNTABILITY AND DEMOCRATIC VULNERABILITY

Beyond selectorate theory, democratic vulnerability to sanctions is rooted in institutional mechanisms of political accountability. Democratic governments operate within systems characterised by competitive elections, independent media, opposition parties, and public scrutiny. Economic deterioration caused by sanctions can therefore generate immediate political consequences for governing leaders. Rising inflation, unemployment, declining living standards, and economic uncertainty often become salient political issues that voters can use to evaluate incumbent performance.

Unlike authoritarian rulers, democratic leaders cannot easily suppress public criticism or prevent opposition groups from mobilising around economic grievances. Media coverage frequently amplifies the visible effects of sanctions, increasing public awareness of economic hardship and creating additional.

pressure on policymakers. As a result, sanctions can transform economic costs into political costs, forcing democratic governments to weigh foreign policy objectives against domestic electoral survival.

However, democratic responsiveness should not be confused with weakness.

Democratic institutions can also provide legitimacy and adaptability, enabling governments to negotiate compromises, implement reforms, and maintain public trust during periods of economic stress. The relationship between democracy and sanctions

vulnerability is therefore complex, reflecting both the constraints and strengths of accountable governance.

INFORMATION CONTROL AND AUTHORITARIAN RESILIENCE

One of the most significant advantages available to authoritarian regimes under sanctions is their ability to control information. Whereas democratic governments generally operate within open information environments, authoritarian states often possess extensive mechanisms for censorship, propaganda, and narrative management. These tools allow governments to shape public perceptions of sanctions and redirect blame away from domestic leadership.

Authoritarian governments frequently portray sanctions as evidence of foreign hostility, external aggression, or attempts to undermine national sovereignty. By framing economic hardship as the consequence of external actors rather than domestic policy failures, regimes can transform sanctions into instruments of political mobilisation. In some cases, sanctions may even strengthen nationalist sentiment and increase public support for incumbent leaders.

Information control also limits the ability of opposition groups to organise around economic grievances. Restrictions on media freedom, civil society organisations, and political dissent reduce the likelihood that economic dissatisfaction will translate into coordinated political pressure. Consequently, the economic costs imposed by sanctions do not automatically generate the political consequences anticipated by sanctioning states.

STATE CAPACITY AND ECONOMIC ADAPTATION

Sanctions resilience depends not only on political institutions but also on the capacity of states to adapt economically. State capacity refers to the ability of governments to mobilise resources, implement policy, enforce regulations, and coordinate responses to external challenges. States possessing strong

Administrative institutions are often better positioned to mitigate the disruptive effects of sanctions.

Governments facing sanctions frequently adopt a range of adaptation strategies.

These include redirecting trade toward alternative partners, developing domestic industries, establishing parallel financial mechanisms, imposing capital controls, and encouraging import substitution. The effectiveness of these strategies varies considerably, but they can significantly reduce the coercive impact of sanctions over time.

Authoritarian regimes often possess greater flexibility in implementing such measures because they face fewer institutional constraints and can rapidly reallocate resources according to political priorities. However, state capacity is not exclusive to

authoritarian systems. Some democratic states also demonstrate remarkable resilience when supported by effective institutions and diversified economies. Consequently, sanctions outcomes are best understood as the product of an interaction between regime characteristics and broader economic capabilities rather than the result of regime type alone.

Together, selectorate dynamics, political accountability, information control, and state capacity provide a comprehensive framework for understanding why sanctions generate divergent outcomes across different political systems. These mechanisms form the basis for the comparative case studies examined in the following sections.

HOW DEMOCRACIES RESPOND: TRANSPARENCY, ACCOUNTABILITY, AND NEGOTIATED COMPLIANCE

1. **Democratic Responses to Economic Sanctions**

Economic sanctions affect democratic states through institutional channels that transform economic costs into political consequences. Because democratic governments depend on public support, sanctions-induced economic hardship often generates pressure for policy adjustment. Rising inflation, declining investment, unemployment, and reduced economic growth can become politically salient issues that voters use to evaluate governing parties. Consequently, democratic leaders frequently face incentives to seek negotiations, compromise, or policy modification when sanctions impose substantial domestic costs.

The presence of independent media further amplifies this dynamic. Journalists, opposition parties, civil society organisations, and interest groups can publicly scrutinise government performance, increasing awareness of economic difficulties and limiting the ability of leaders to redirect blame. Electoral competition ensures that political opponents can capitalise on public dissatisfaction, transforming economic grievances into political challenges.

However, democratic responsiveness should not be interpreted solely as vulnerability. Democratic institutions may also enhance resilience by fostering legitimacy, transparency, and public trust. Citizens are often more willing to tolerate temporary hardship when governments maintain credibility and clearly communicate policy objectives. As a result, democratic responses to sanctions vary considerably depending on public opinion, institutional strength, and the perceived legitimacy of both domestic and international actors.

Authoritarian governments confront sanctions within a fundamentally different political environment. Unlike democratic leaders, authoritarian rulers typically face fewer institutional mechanisms of accountability and can therefore absorb economic pressure

without experiencing immediate political consequences. Economic hardship does not automatically translate into organised political opposition because opportunities for dissent are frequently restricted through censorship, surveillance, and political repression.

A defining characteristic of authoritarian resilience is the ability of regimes to selectively distribute economic costs. Through control over state resources, governments can shield politically important groups such as military leaders, party officials, and economic elites while transferring a disproportionate share of the burden onto the broader population. This reduces the likelihood that sanctions will undermine the loyalty of those whose support is essential for regime survival.

Authoritarian governments also possess advantages in information management.

By controlling media narratives and restricting access to alternative sources of information, regimes can portray sanctions as acts of foreign aggression rather than consequences of domestic policy choices. Such narratives often strengthen nationalist sentiment and encourage public support for resistance against external pressure. Rather than weakening authoritarian governments, sanctions may sometimes contribute to a "rally-around-the-flag" effect that reinforces regime legitimacy.

LIMITATIONS OF THE DEMOCRACY-AUTOCRACY FRAMEWORK

Although regime type provides a valuable framework for understanding sanctions outcomes, it should not be treated as a complete explanation. Democracies and

Autocracies are not internally uniform categories. Significant variation exists within both groups, influencing their capacity to withstand external pressure.

Among authoritarian systems, single-party states and military regimes often demonstrate greater resilience than personalist dictatorships because they possess stronger institutions and more stable mechanisms for elite coordination. Similarly, democratic states vary in levels of political polarisation, institutional strength, economic diversification, and public trust. These differences can substantially affect how sanctions are experienced and managed.

Furthermore, factors such as economic size, resource endowments, geographic location, and access to alternative trading partners interact with regime characteristics to shape outcomes. A resource-rich authoritarian state with strong external partnerships may prove far more resilient than a small, isolated authoritarian regime. Likewise, some democratic states may successfully withstand sanctions when supported by effective institutions and broad public legitimacy.

Consequently, regime type should be understood as one of the most significant determinants of sanctions resilience rather than the sole determinant. The following case studies examine how these theoretical mechanisms operate in practice through the experiences of Russia, Iran, and South Africa.

CASE STUDY I: RUSSIA AND THE LIMITS OF ECONOMIC COERCION

The sanctions imposed on Russia following its February 2022 invasion of Ukraine constitute one of the most extensive and coordinated economic coercion campaigns in modern history. The United States, the European Union, the United Kingdom, Canada, Japan, and numerous other states implemented a broad array of measures designed to isolate Russia from the global economy. These measures included restrictions on major financial institutions, the freezing of foreign reserves, export controls on advanced technologies, limitations on access to international payment systems, and sanctions targeting political elites, state-owned enterprises, and key sectors of the Russian economy. Many policymakers and analysts initially predicted that the scale of these restrictions would generate severe economic disruption and compel significant changes in Russian policy.

The immediate economic effects were substantial. Following the announcement of sanctions, the Russian rouble experienced sharp volatility, foreign investment declined significantly, and numerous multinational corporations suspended or terminated operations within Russia. Restrictions on access to Western technology disrupted sectors dependent upon imported components, while financial sanctions complicated matters.

international transactions and investment flows. The freezing of a significant portion of Russia's foreign reserves represented an unprecedented effort to constrain the state's ability to stabilise its economy and finance its activities.

Despite these challenges, Russia demonstrated a level of resilience that exceeded many initial expectations. Rather than experiencing immediate economic collapse, the Russian government implemented a series of adaptation measures consistent with the theoretical mechanisms discussed earlier. Capital controls were introduced to stabilise the currency, financial institutions were restructured to accommodate sanctions restrictions, and state resources were redirected towards sectors deemed strategically important. These measures reflected the state's capacity to rapidly coordinate economic responses and minimise short-term instability.

A crucial factor in Russia's resilience was its ability to maintain access to alternative economic partners. Although trade with many Western countries declined sharply, Russia expanded economic relationships with states that did not fully participate in the sanctions regime. Energy exports, particularly oil and natural gas, continued to generate substantial revenue, while trade relationships with countries such as China and India provided alternative markets for Russian goods. This diversification reduced the effectiveness of efforts to economically isolate the country and demonstrated the importance of international economic networks in shaping sanctions outcomes.

Political factors proved equally significant. Russia's authoritarian political structure limited the extent to which economic hardship translated into direct political pressure on the governing elite. Consistent with selectorate theory, the regime prioritised the protection of politically important groups while distributing costs across broader society. At the same time, extensive control over media and information environments enabled authorities to shape public perceptions of sanctions. Economic difficulties were frequently presented as the result of Western hostility rather than domestic policy choices, reinforcing nationalist narratives and reducing the likelihood that economic grievances would evolve into organised political opposition.

The Russian case illustrates both the strengths and limitations of sanctions as instruments of coercion. The sanctions undoubtedly imposed high economic costs, restricted access to technology, and complicated long-term economic development. However, they did not produce the rapid policy reversal or political destabilisation that many observers initially anticipated. Instead, Russia's experience demonstrates how authoritarian governance, state capacity, information control, and access to alternative economic partnerships can collectively enhance sanctions resilience.

Rather than proving that sanctions are ineffective, the Russian case suggests that economic pressure alone may be insufficient to compel major policy changes when targeted states possess the institutional and geopolitical resources necessary to adapt. The case, therefore, provides strong support for the argument that regime characteristics significantly influence how sanctions are experienced, managed, and ultimately resisted.

CASE STUDY II: IRAN AND LONG-TERM SANCTIONS ADAPTATION

Iran represents one of the most significant examples of long-term sanctions resilience in the modern international system. Since the 1979 Islamic Revolution, the country has been subjected to multiple rounds of American, European, and multilateral sanctions targeting its financial sector, energy exports, military activities, and nuclear programme. These sanctions have sought to alter Iranian behaviour across a range of issues, including nuclear development, regional influence, and support for armed groups. Despite the breadth and duration of these measures, Iran has repeatedly demonstrated an ability to endure substantial economic pressure while maintaining many of its core strategic policies.

The most intensive sanctions period emerged during disputes surrounding Iran's nuclear programme. Beginning in the mid-2000s, sanctions increasingly targeted Iran's banking system, oil exports, and access to international financial networks. Restrictions on energy exports significantly reduced government revenues, while financial sanctions complicated international trade and investment. The Iranian economy experienced high inflation, currency depreciation, declining foreign investment, and periodic recessions.

These developments imposed substantial costs on both the state and the wider population.

Nevertheless, sanctions did not produce regime collapse or immediate policy capitulation. Instead, Iran adopted a range of adaptation strategies that reduced the

effectiveness of external pressure. Successive governments promoted what officials described as a "resistance economy," emphasising domestic production, reduced dependence on foreign imports, and the expansion of alternative economic partnerships. Trade relations with countries such as China, Russia, and several regional actors helped compensate for restrictions imposed by Western states. Informal economic networks and sanctions-evasion mechanisms further enabled the continuation of international commerce despite significant obstacles.

Political institutions also contributed to Iran's resilience. Although Iran contains electoral elements, ultimate authority remains concentrated within unelected institutions, including the Supreme Leader and key security organisations. This hybrid structure limits the extent to which economic dissatisfaction can directly threaten regime survival.

Consistent with selectorate theory, political leaders have often prioritised the protection of strategically important constituencies while distributing economic burdens across broader segments of society. As a result, severe economic hardship has not necessarily translated into the political pressure anticipated by sanctioning states.

Information control has similarly played an important role. Iranian authorities have frequently portrayed sanctions as evidence of foreign hostility and attempts to undermine national sovereignty. Such narratives frame economic hardship as a consequence of

external aggression rather than solely domestic policy choices, helping to sustain nationalist sentiment and reduce the political effectiveness of sanctions. Although public dissatisfaction and protests have periodically emerged, these movements have not fundamentally altered the structure of political power.

At the same time, the Iranian case demonstrates that sanctions are not without effect. Economic growth has been constrained, inflation has remained persistently high, and access to advanced technology and foreign investment has been significantly restricted. Sanctions contributed to the conditions that ultimately produced negotiations leading to the Joint Comprehensive Plan of Action (JCPOA) in 2015, illustrating that economic pressure can influence policy calculations even when it fails to achieve broader political transformation.

The Iranian experience, therefore, highlights both the possibilities and limitations of economic coercion. Sanctions imposed substantial economic costs and influenced specific policy decisions, yet they did not fundamentally alter the political structure of the state or compel comprehensive compliance with external demands. Iran's ability to combine institutional resilience, economic adaptation, information management, and alternative international partnerships provides strong support for the argument that authoritarian and semi-authoritarian systems often possess significant advantages in resisting prolonged sanctions pressure.

CASE STUDY III: SOUTH AFRICA AND THE CONDITIONAL SUCCESS OF SANCTIONS

South Africa's experience during the final decades of apartheid provides one of the most frequently cited examples of sanctions contributing to political change.

Throughout the 1970s and 1980s, growing international opposition to apartheid led numerous states, international organisations, corporations, and civil society groups to impose economic, financial, cultural, and diplomatic restrictions on the South African government. Unlike many contemporary sanctions campaigns, these measures were accompanied by extensive international advocacy efforts and increasing domestic resistance, creating a multifaceted pressure environment that ultimately contributed to the dismantling of apartheid.

The sanctions campaign intensified during the 1980s as international condemnation of apartheid expanded. Major trading partners imposed restrictions on investment, financial transactions, arms sales, and various forms of economic cooperation. International corporations withdrew investments, foreign banks became increasingly reluctant to extend credit, and South Africa faced growing difficulties accessing international capital markets. Although the South African economy did not collapse, sanctions significantly increased economic costs and contributed to declining investor confidence, slower economic growth, and mounting financial pressures.

Unlike the authoritarian systems examined in the previous case studies, South Africa's political structure contained important features that increased its vulnerability to external pressure. While apartheid denied political rights to the majority of the population, the regime nevertheless depended heavily on cooperation from business leaders, economic institutions, and segments of the white electorate whose economic interests were increasingly threatened by international isolation. Economic deterioration, therefore, generated pressure not only from excluded populations but also from influential groups within the political and economic establishment.

The interaction between sanctions and domestic political dynamics proved particularly significant. Internal resistance movements, labour organisations, community groups, and anti-apartheid activists had already created substantial challenges for the government. Sanctions did not operate independently of these forces; rather, they amplified existing economic and political pressures. As the costs of maintaining apartheid increased, influential actors within South African society began to question the long-term sustainability of the system. Economic concerns became increasingly intertwined with broader political debates regarding reform and transition.

The South African case also illustrates the importance of international coordination. Sanctions were reinforced by widespread diplomatic pressure, global media attention, and transnational advocacy networks that maintained international focus on apartheid. This broad coalition reduced opportunities for the regime to circumvent restrictions and increased the reputational costs associated with continued resistance. The relative effectiveness of sanctions in South Africa, therefore, reflected not only domestic political conditions but also the unusually high degree of international consensus supporting the sanctions effort.

However, attributing the end of apartheid solely to sanctions would oversimplify a complex historical process. Political negotiations, internal resistance, changing geopolitical conditions, economic transformations, and leadership decisions all played essential roles in shaping the transition. Sanctions functioned as one component of a broader constellation of pressures rather than as an independent or decisive cause.

Their contribution was significant, but it operated in combination with numerous domestic and international factors.

South Africa, therefore, serves as an important counterpoint to the experiences of Russia and Iran. Whereas those cases demonstrate how authoritarian systems can absorb and manage external economic pressure, South Africa illustrates the circumstances under which sanctions can contribute to meaningful political change. The case suggests that sanctions are most likely to influence outcomes when economic costs affect politically influential groups, domestic opposition movements remain active, and international actors maintain broad and sustained coordination. Consequently, sanctions effectiveness depends not only on the severity of economic pressure but also on the political environment through which that pressure is transmitted.

COMPARATIVE ANALYSIS

1. **Political Accountability and Elite Pressure**

One of the clearest differences between the cases concerns the relationship between economic hardship and political accountability. Sanctions are designed to impose economic costs with the expectation that these costs will generate political pressure on decision-makers. However, the extent to which economic pain translates into political consequences depends heavily on the structure of political institutions.

In both Russia and Iran, political leaders were largely insulated from direct public accountability. Economic difficulties resulting from sanctions undoubtedly affected ordinary citizens, yet these costs did not automatically threaten regime survival.

Consistent with selectorate theory, both governments prioritised the protection of strategically important elites whose support remained essential to maintaining political stability. As long as these groups continued to benefit from state resources and remained loyal to the regime, sanctions struggled to generate sufficient pressure for major policy change.

South Africa presented a different dynamic. While the apartheid regime was not a democracy, it nevertheless depended upon the cooperation of influential economic actors, business leaders, and segments of the white electorate. As international isolation intensified and economic costs increased, these groups experienced growing pressure to reconsider the sustainability of apartheid. The South African case, therefore, demonstrates that sanctions become more effective when economic costs directly affect politically influential constituencies rather than being concentrated solely on the general population.

Taken together, the cases suggest that sanctions are most effective when they reach the groups upon whom political leaders depend for survival. Economic suffering alone does not guarantee political change; what matters is who bears the costs and how those costs affect the regime's core support base.

INFORMATION CONTROL AND PUBLIC PERCEPTION

The comparative evidence also highlights the critical role of information control in shaping sanctions outcomes. Economic pressure can only influence political behaviour if citizens attribute economic hardship to government policies and possess the capacity to express dissatisfaction effectively.

Russia and Iran both demonstrated substantial ability to shape public narratives surrounding sanctions. Government-controlled media frequently portrayed sanctions as evidence of foreign hostility, external aggression, or attempts to undermine national interests.

sovereignty. By framing economic difficulties as the result of actions taken by external actors rather than domestic policy failures, both regimes reduced the likelihood that public dissatisfaction would be directed toward political leadership.

In contrast, South Africa faced a far more challenging information environment.

Domestic resistance movements, international media coverage, and global anti-apartheid activism ensured that sanctions remained closely connected to criticism of apartheid itself. The government found it increasingly difficult to control narratives surrounding international isolation and economic decline. Consequently, sanctions contributed to growing political debate regarding the legitimacy and sustainability of the existing system.

The comparison demonstrates that sanctions are more likely to influence political outcomes when governments possess limited capacity to control information and shape public perceptions. Information control can significantly weaken the political consequences of economic hardship by redirecting blame and reinforcing regime legitimacy.

ECONOMIC ADAPTATION AND STATE CAPACITY

A third factor distinguishing the cases is the ability of states to adapt economically to external pressure. Sanctions rarely operate in static environments; targeted governments actively seek methods of mitigating their effects through policy adjustments, alternative partnerships, and institutional reforms.

Russia responded to sanctions through a combination of capital controls, trade redirection, financial restructuring, and continued reliance on energy exports. Access to alternative markets, particularly in Asia, helped offset losses resulting from reduced economic engagement with Western states. Similarly, Iran developed extensive

adaptation mechanisms over decades of sanctions, including import substitution, informal trade networks, and closer economic relationships with non-Western partners.

South Africa, while possessing a relatively advanced economy, faced greater challenges in offsetting the effects of international isolation. Restrictions on investment and finance, combined with declining investor confidence, created pressures that proved increasingly difficult to manage. Unlike Russia and Iran, South Africa had fewer opportunities to rely on alternative geopolitical partnerships capable of fully compensating for the costs imposed by sanctions.

The evidence, therefore, suggests that state capacity and economic adaptability play crucial roles in determining sanctions resilience. Sanctions are often most effective during the initial stages of implementation, before targeted states develop mechanisms for adaptation and circumvention.

BEYOND REGIME TYPE: ADDITIONAL DETERMINANTS

Although the analysis strongly supports the importance of regime type, the case studies also demonstrate that political institutions alone cannot explain sanctions outcomes. Several additional variables significantly influenced the effectiveness of sanctions across the three cases.

Economic size and resource endowments were particularly important. Russia benefited from substantial energy exports that continued to generate revenue despite international restrictions. Iran similarly relied upon strategic resources and regional partnerships that reduced the severity of economic isolation. South Africa's integration into global financial markets, by contrast, increased its vulnerability to coordinated international pressure.

International coordination also affected outcomes. Sanctions tend to become more effective when participation among sanctioning states is broad, and opportunities for circumvention are limited. The anti-apartheid movement achieved an unusually high degree of international consensus, whereas sanctions against Russia and Iran have been undermined to varying degrees by the continued engagement of non-participating states.

These observations reinforce the conclusion that sanctions effectiveness emerges from the interaction between political institutions, economic conditions, and international dynamics rather than from any single factor in isolation.

KEY FINDINGS

The comparative analysis yields three principal findings. First, sanctions are most likely to influence political outcomes when economic costs reach politically influential groups whose support is necessary for regime survival. Second, information control significantly affects whether economic hardship is translated into political pressure or

redirected toward external actors. Third, strong state capacity and access to alternative economic partnerships enhance sanctions resilience by reducing the long-term effectiveness of economic coercion.

Taken together, these findings support the central argument of this article: regime type is among the most important determinants of sanctions resilience because it shapes how economic costs are distributed, interpreted, and managed. However, regime type does not operate independently. Its effects are mediated by broader economic and geopolitical factors that can either strengthen or weaken the coercive impact of sanctions. Understanding this interaction is essential for explaining why some sanctions campaigns succeed while others fail despite imposing substantial economic costs.

POLICY IMPLICATIONS

1. **Designing More Effective Sanctions**

The findings of this study suggest that policymakers should reconsider the assumption that economic pain automatically generates political compliance. While sanctions can impose substantial economic costs, their effectiveness depends on how those costs are distributed within the target state. In authoritarian systems, governments often possess the ability to shield politically important elites while transferring a disproportionate share of hardship onto the general population. Consequently, broad economy-wide sanctions may generate severe economic disruption without significantly altering the calculations of political leaders.

To address this challenge, policymakers should place greater emphasis on targeted or "smart" sanctions directed at individuals, institutions, and economic sectors that are directly connected to regime survival. Measures targeting political elites, military organisations, state-owned enterprises, and strategic industries may increase pressure on decision-makers while reducing unintended humanitarian consequences for ordinary citizens. By focusing on the actors most capable of influencing government policy, sanctions are more likely to generate meaningful political leverage.

Furthermore, sanctions should be designed with a clear understanding of the domestic political structure of the target state. Strategies that may prove effective against democratic governments may be considerably less effective against authoritarian systems. A regime-specific approach is therefore likely to produce better outcomes than a one-size-fits-all sanctions policy.

THE IMPORTANCE OF INTERNATIONAL COORDINATION

The experiences of Russia and Iran demonstrate that sanctions are significantly less effective when targeted states retain access to alternative economic partners. Both countries mitigated the effects of sanctions through trade diversification, alternative

financial arrangements, and economic cooperation with states that did not fully participate in sanctions regimes. These adaptation strategies reduced the degree of economic isolation intended by sanctioning actors.

As a result, international coordination should be viewed as a critical component of sanctions effectiveness. Broad multilateral participation limits opportunities for circumvention, increases economic pressure, and enhances the credibility of sanctions campaigns. By contrast, unilateral or narrowly coordinated sanctions often leave significant loopholes that targeted states can exploit.

Effective enforcement is equally important. Even comprehensive sanctions can lose effectiveness when implementation is inconsistent or when states fail to adequately monitor compliance. Policymakers should therefore prioritise international cooperation not only during the design phase of sanctions but also throughout their implementation and enforcement.

SANCTIONS AS PART OF A BROADER STRATEGY

The South African case illustrates that sanctions are rarely most effective when used in isolation. Economic pressure contributed to political change, not because it operated alone, but because it reinforced domestic opposition, diplomatic pressure, international advocacy efforts, and broader political transformations. This suggests that sanctions should be viewed as one component of a wider strategy rather than as a standalone solution.

Policymakers should therefore integrate sanctions with diplomatic engagement, negotiation frameworks, and clearly defined pathways for compliance. Target states are more likely to alter their behaviour when sanctions are accompanied by credible incentives and realistic opportunities for policy adjustment. Without such mechanisms, sanctions risk becoming instruments of punishment rather than tools of political influence.

Finally, governments should adopt realistic expectations regarding what sanctions can achieve. Sanctions may successfully signal international condemnation, constrain economic activity, or increase the costs of particular policies without necessarily producing immediate political concessions. Evaluating sanctions solely on the basis of regime change or complete policy reversal risks overlooking more limited but nevertheless meaningful outcomes. A nuanced understanding of sanctions effectiveness can contribute to more effective policy design and more accurate assessments of success.

The evidence presented in this article, therefore, suggests that successful sanctions strategies require more than economic pressure alone. They require careful consideration of political institutions, elite incentives, information environments, state capacity, and international coordination. Policymakers who account for these factors are more likely to design sanctions regimes capable of generating meaningful political influence rather than merely imposing economic costs.

LIMITATIONS

While this study provides evidence that regime type significantly influences sanctions resilience, several limitations should be acknowledged. First, the analysis relies on a relatively small number of case studies. Although Russia, Iran, and South Africa provide valuable comparative insights, they cannot fully capture the diversity of political systems and sanctions experiences observed throughout the international system. Future research incorporating a larger sample of cases could provide a more comprehensive assessment of the relationship between regime type and sanctions outcomes.

Second, the Russia case remains ongoing. Unlike the historical experiences of Iran and South Africa, the long-term political and economic consequences of the sanctions imposed after the 2022 invasion of Ukraine are still unfolding. Any assessment of Russian sanctions resilience must therefore be regarded as provisional and subject to revision as additional evidence becomes available.

Third, the study focuses primarily on political institutions while giving comparatively less attention to structural factors such as geography, resource endowments, economic size, and regional power dynamics. These variables can significantly influence sanctions outcomes and, in some cases, may interact with regime type in ways that are difficult to isolate. For example, the resilience demonstrated by Russia is partly linked to its energy resources and access to alternative trading partners, factors that may not be available to other authoritarian states.

Finally, the effectiveness of sanctions is inherently difficult to measure. Sanctions often pursue multiple objectives simultaneously, including deterrence, signalling, punishment, and behavioural change. Determining whether a sanctions regime has succeeded, therefore, depends partly on how success itself is defined. Future research could benefit from examining different measures of sanctions effectiveness and exploring how these measures vary across political systems.

While these limitations should be considered when interpreting the findings, they do not diminish the value of examining sanctions through the lens of political institutions. Rather, they highlight the need for further research into the interaction between regime type, economic structure, and international dynamics. Such research may contribute to a more comprehensive understanding of why some sanctions campaigns generate meaningful political change while others fail to achieve their intended objectives.

CONCLUSION

Economic sanctions have become one of the most frequently employed instruments of international statecraft, occupying a position between diplomacy and military intervention as a means of influencing state behaviour. Yet despite their widespread use, questions regarding their effectiveness continue to be raised.

considerable scholarly and policy debate. This article has sought to contribute to that debate by examining how regime type influences a state's capacity to withstand and adapt to economic sanctions.

The analysis demonstrates that sanctions do not operate solely through economic mechanisms. Rather, their effectiveness depends upon the political institutions through which economic costs are distributed, interpreted, and translated into political consequences. Drawing upon selectorate theory, political accountability, information control, and state capacity, this study has argued that authoritarian regimes often possess structural advantages that enable them to resist sanctions pressure more effectively than democratic or semi-democratic systems. Narrower accountability structures, greater control over information environments, and enhanced flexibility in resource allocation frequently allow authoritarian governments to absorb economic hardship without facing the political consequences anticipated by sanctioning states.

The comparative case studies provide substantial support for this argument. Russia's response to the extensive sanctions imposed following its 2022 invasion of Ukraine illustrates how state capacity, information control, and alternative economic partnerships can mitigate the impact of even unprecedented economic restrictions. Iran's decades-long experience under sanctions demonstrates the ability of a state to develop adaptation mechanisms that reduce the effectiveness of prolonged economic pressure. In contrast, South Africa highlights circumstances under which sanctions can contribute to meaningful political change, particularly when economic costs affect influential constituencies, domestic opposition remains active, and international coordination is sustained.

At the same time, the findings caution against viewing regime type as a complete explanation for sanctions outcomes. Economic structure, natural resource endowments, geographic position, international partnerships, and the degree of multilateral cooperation all influence the effectiveness of sanctions. The experiences of Russia, Iran, and South Africa demonstrate that sanctions outcomes emerge from the interaction of political, economic, and geopolitical factors rather than from any single variable alone.

The broader implication of this study is that sanctions should not be evaluated solely according to the economic damage they inflict. Economic hardship does not automatically generate political change. Instead, the success or failure of sanctions depends on whether economic costs reach politically influential actors, whether governments can control public narratives surrounding those costs, and whether states possess the institutional capacity to adapt to external pressure. Understanding these mechanisms is essential for explaining why some sanctions campaigns achieve their objectives while others fail despite imposing substantial economic burdens.

As economic sanctions continue to play a central role in international relations, policymakers and scholars alike must move beyond simplistic assumptions regarding their effectiveness. The central lesson of this study is not that sanctions work or fail.

universally, but that their outcomes are shaped by the political systems they seek to influence. Appreciating this reality is crucial for designing more effective sanctions

regimes and for developing a more nuanced understanding of economic coercion in contemporary world politics.

REFERENCES

Allen, S. H. (2005). The determinants of economic sanctions success and failure.

International Interactions, 31 (2), 117-138.

Bueno de Mesquita, B., Smith, A., Siverson, R. M., & Morrow, J. D. (2003). The logic of political survival. MIT Press.

Chatham House. (2025). Understanding and improving sanctions today: Some real and imagined concerns. Chatham House.

Congressional Research Service. (2024). Russia--North Korea relations (IF12760).

Congressional Research Service.

Crawford, N. C., & Klotz, A. (Eds.). (1999). How sanctions work: Lessons from South Africa. St. Martin's Press.

Elgar Online. (2024). Pressure proofing: How authoritarian regimes respond to sanctions. In the Research handbook on authoritarianism. Edward Elgar Publishing.

Escribà-Folch, A. (2012). Authoritarian responses to foreign pressure: Spending, repression, and sanctions. Comparative Political Studies, 45 (6), 683-713. <https://doi.org/10.1177/0010414011427888>

Freedom House. (2022). Freedom in the world 2022: The global expansion of authoritarian rule. Freedom House.

Grauvogel, J., & von Soest, C. (2014). Claims to legitimacy count: Why sanctions fail to instigate democratisation in authoritarian regimes. European Journal of Political Research, 53 (4), 635-653. <https://doi.org/10.1111/1475-6765.12065>

Hufbauer, G. C., Schott, J. J., Elliott, K. A., & Oegg, B. (2007). Economic sanctions reconsidered (3rd ed.). Peterson Institute for International Economics.

International Monetary Fund. (2023). Russian economy amid sanctions: Recent developments and outlook. International Monetary Fund.

Klotz, A. (1995). Norms in international relations: The struggle against apartheid. Cornell University Press.

RAND Corporation. (2023). *Developing bad habits: What Russia might learn from Iran's sanctions evasion*. RAND Corporation.

United Nations. (1945). *Charter of the United Nations* (Chapter VII, Articles 39-42).

United Nations.

van Bergeijk, P. A. G. (2021). *Research handbook on economic sanctions*. Edward Elgar Publishing.

World Bank. (2023). *Russia economic update: Adjustment and resilience under sanctions*. World Bank.

World Development. (2023). Sanctions effectiveness, development and regime type: Are aid suspensions and economic sanctions alike? *World Development*.

European Political Science. (2023). Rally-around-the-flag effects in the Russian-Ukrainian war. *European Political Science*.

Defence and Peace Economics. (2019). Autocracies and economic sanctions: The divergent impact of authoritarian regime type on sanctions success. *Defence and Peace Economics, 30* (3), 286-302.

Journal of Conflict and Security Law. (2022). Legality of unilateral extraterritorial sanctions under international law. *Journal of Conflict and Security Law, 27* (1), 53-76.

Journal of International Law & Human Rights. (2024). The legality of unilateral economic sanctions in international law. *Journal of International Law & Human Rights, 3* (1), 1-22.

PAPER 08 / 08

Civil War of Myanmar

A Study of Political Turmoil and Strategic Friction Between Major External Actors

Veer P. Singh · Christu Jyoti Convent Senior Secondary School, Baraut

ABSTRACT

After 2021 Myanmar coup d'état, Myanmar has been plunged into a deep civil war, blocking various pathways and opening some for its neighbors. There has been an influx of refugees in neighboring countries, disruption of multi-million dollar worth of trade corridors, and a new wave of geopolitical agendas.

Beijing views Myanmar as a source of rare earth minerals, but also a source of deep concern securing economic corridors to Indian Ocean. Similarly, New Delhi sees Myanmar as a way to open its north eastern states to sea routes, but also concerns over the refugee spillovers from Myanmar, like Rohingyas, to protect internal security.

This article argues that the fragmentation of Myanmar has create a strategic stalemate for its neighbors where external powers can neither afford to ignore the crisis nor successfully intervene to resolve it. This civil war opens a great skirmish between the neighboring powers over geographical strategic locations and financial control over Naypyidaw.

Keywords: Myanmar coup d'état, rare earth minerals, India, China, Rohingyas, Naypidaw

CIVIL WAR OF MYANMAR

Myanmar as a nation has been in the grip of the power of its internal military forces since the very beginning of its independence. Though, there was some time period of democratic power, after the coup of 2021, the political status of Myanmar has changed drastically. The nation has been going through an ongoing civil war that includes the military, the ousted government, many ethnic guerrilla fighters, the civilians, and the neighboring nations, looking for their own needs and agendas.

China and India have been constantly involved with the Myanmar civil war.

- China is a main player in the Myanmar civil war. China gives somewhat near \$3 billion to the junta as development funds and aids. China covets rare earth minerals found in Myanmar, the junta has permitted the expansion of unregulated mining operations in Kachin State, many of which are supplying Chinese entities (Naing & Sae-phoo, 2025). But, china also has to deal with kachin independence army, which disrupts the supply of minerals. This is mainly driven by the fact that china wants to secure its energy supplies, like any other nation would have done,

- India as a neighbor of Myanmar has to deal with a softer kind of diplomacy with Myanmar. India has developed the Kaladan Multi-Modal Transit Transport Facility to connect its north eastern states with the sea. The corridor passes through the Rakhine and Chin regions of Myanmar, controlled by Arakan Army. So, India has to go through a very soft and shrewd type of diplomacy. India shares highly porous borders with Myanmar, which are easily infiltrated by illegal immigrants. As found in Singh, 2026, Myanmar serves as a an illegal opiate supplier and meth hub through these said borders. So, it becomes impossible for India to stay in a neutral position on the topic of civil war.

This paper examines the distinct geopolitical motives and driving complications that shape the responses of China and India towards the Myanmar civil war. Ultimately, this article states that because these external forces are driven by fundamentally irreconcilable strategic interests, a unified diplomatic effort to stop the civil war remains impossible. As a result, the conflicting priorities of its neighbors inadvertently sustain a protracted stalemate.

CHINA

CHINA'S STRATEGIC USE OF MYANMAR'S JUNTA

In alignment with its core, Beijing's foreign policy towards the junta is driven strictly by domestic security and economic imperatives, viewing the military regime as a rational partner necessary for safeguarding its economic and geopolitical interests in the region. China shares its southern border with Myanmar, which spans roughly 2,204km ("Ministry of Commerce of Myanmar", n.d.). After the coup of 2021, there has been a shift in how china deals with the civil war. Beijing has cultivated a highly pragmatic relationship with Naypidaw, extending financial aid in exchange for exclusive concessions to extract rare earth minerals and strategic metals.

Beijing has had a very different response to the Rohingya genocide and displacement by the Tatmadaw (Myanmar's military). Beijing supports the atrocities, stating that these are Myanmar's domestic affairs and that it is fighting "Terrorism". There is an underlying energy need that drives this stance. China covets the vast reserves of oil and gas in the Rakhine state of Myanmar, home to the Rohingyas (Huq, 2026). This supposedly points to the fact of china having a neutral stand on the Rohingya massacre.

Central to this is the China--Myanmar Economic Corridor which links Yunnan to the Bay of Bengal through pipelines, roads, railways and port projects. The Kyaukphyu deep-sea port in Rakhine State in particular offers a potential direct gateway to the Indian Ocean, reducing China's reliance on the Malacca Strait (Ko, 2026). Consequently, Beijing's diplomatic framework intentionally prioritizes structural stability over humanitarian concerns within Myanmar (Haacke, 2019). Historically, securing resource concessions and infrastructural access proves far more streamlined under a dependent military regime than under a democratic administration accountable to public scrutiny and environmental regulations (Kurlantzick, 2021)

PROBLEMS ARISING FROM THE CIVIL WAR

Aside from the advantages china gets from the military rule and fragmentation of Myanmar, it also possesses some deep regional troubles due to uneven governing bodies of different regions. On October 27, 2023, the Three Brotherhood Alliance (3BTA), composed of the Arakan Army, Myanmar National Democratic Alliance Army, and Ta'ang National Liberation Army and numbering somewhere around 15,000 fighters, launched Operation 1027, an anti-SAC offense in Shan State. The armed conflict captured significant territory bordering China (Sun, 2024)

These kinds of armed conflicts between junta and anti junta forces disrupt china's mining projects, ports in Myanmar, and trading corridors. Beijing maintains suboptimal relationships with these kinds of Ethnic Armed Organizations (EAOs). These relations are important if china wants to protect its valuable assets in region controlled by non-state armed actors.

INDIA

INDIA'S DILEMMA BETWEEN PEACE AND SECURITY

India strives for peace and democracy on the international stage. As the initiator of Non-Aligned Movement, India does not side with any of the power blocks in the world. But when it comes to Myanmar, it becomes a difficult choice for India. India, as a nation, is currently between a dilemma between its own national security, and managing peace and security of Myanmar because of its international stance (Kumbhar, 2023).

The north-eastern states of India are in constant troubles because of the escalating conflict in Myanmar. India shares its far north-eastern border with Myanmar, a highly volatile zone of ethnical clashes and guerrilla fighters. The conflict in Myanmar poses a threat. For India, Myanmar provides land and maritime connectivity as a gateway to Southeast Asia, highlighting the significance of the cooperation. However, this unrest in Myanmar has hampered India's economic and connectivity initiatives. The same has also impacted the completion of India's Developmental Assistance (IDA) in Myanmar; the \"Kaladan Multi-Modal Transit Transport Project,\" which involves road, rail and marine transportation from Sittwe Port in Myanmar to the Northeastern region; and the India-Myanmar-Thailand trilateral highway (IMT), to name a few (Lamkang, 2025).

Far from maintaining passive neutrality, New Delhi's engagement with Tatmadaw is a vital domestic security imperative, dictated by the need to secure its highly volatile northeast frontier and deny cross border sanctuaries to active insurgent groups like the NSCN-K and ULFA (Singh, 2024).

ROHINGYA CRISIS

The Rohingya are a predominantly Muslim Indo-Aryan ethnic minority indigenous to Rakhine State, Myanmar. Denied citizenship under Myanmar\'s 1982 law, they form the world\'s largest stateless population. These people therefore have been living in overcrowded refugee camps in neighboring countries like Bangladesh and India.

After the coup, the persecution of Rohingyas has been increasing because of a state of anarchy and ethnical persecution. Indian government stands firmly that the Rohingya poses a threat on national security and resources. On 4 October 2018, a three judge bench of the Supreme Court of India refused to stay the deportation of seven Rohingya refugees from India to Myanmar (Samtani, 2018).

States in India governed by the Hindu nationalist Bharatiya Janata Party (BJP) initiated a campaign to expel Rohingya and Bengali-speaking Muslims for being "illegal immigrants." Those expelled to Bangladesh included at least 192 Rohingya refugees despite being registered with the United Nations refugee agency (UNHCR). The authorities also put 40 Rohingya refugees on a ship near the Myanmar coast and forced them to swim ashore. Dozens more have fled to Bangladesh to avoid the crackdown. ("India: scores of Rohingya refugees expelled", 2025).

Consequently, the protracted civil conflict in Myanmar imposes severe human security and administrative drain on India. New Delhi must continuously mitigate the transnational fallout of the crisis, which manifests as illicit trafficking corridors along the porous frontier and complex migratory waves into its border states. This forces India to adopt a dual-track diplomacy, maintaining fluid transactions with both the Tatmadaw and various anti-junta coalitions. Rather than a passive stance, this cautious approach represents a classic strategic compromise: balancing India's normative, democratic state identity with the realist imperatives of maintaining border stability and protecting critical infrastructure assets from regional disruptions.

CONCLUSION

Ultimately, this paper shifts the analytical focus from the internal dynamics of Myanmar's civil war to the external calculated interests that freeze the conflict in place. The structural divergence between china's infrastructure driven relations with the junta and India's security driven, dual track diplomacy toward Myanmar's different powers ensures that no singular regional authority has the incentive or will to enforce a peace plan for Myanmar. The regional powers have effectively adapted to a highly fragmented Myanmar, treating territorial volatility as a risk to be insulated against rather as a problem that needs to be solved.

This structural deadlock answers the crucial 'so what?' of the crisis: as long as Myanmar's internal governance remains a battleground for bilateral hegemony and border containment, the state will continue to function as a fractured buffer zone. The international community therefore must view the Myanmar civil war not as a case of localized anarchy, but as a permanently managed geopolitical stalemate where systematic external competing interests actively prevents a unified resolution.

REFERENCES

Aung Thura Ko Ko. (2026). China deepens its foothold in post-coup Myanmar. Cornell University. <https://www.cornell.edu/>

Association of Southeast Asian Nations (ASEAN). (2021, April 24). Chairman's statement on the Leader\'s Meeting. ASEAN Secretariat.

Epstein, A. K. (2024, October 2). ASEAN still torn over security challenges (A. K. Epstein, Illus.).

Haacke, J. (2019). China's role in the Myanmar crisis: Realist diplomacy and the limits of structural influence. International Affairs, 95(5), 1011-1031.

Human Rights Watch. (2025, August 28). India: Scores of Rohingya refugees expelled; Hundreds more arbitrarily detained, threatened with deportation. <https://www.hrw.org/news/2025/08/28/india-scores-rohingya-refugees-expelled>

Huq, S. (n.d.). How China-Myanmar relations impact the Rohingya crisis. <https://refugeeresearchonline.org/how-china-myanmar-relations-impact-the-rohingya-crisis/>

Kumbhar, S. (2023). Peace and conflict resolution: The Indian perspectives. Journal of National Development, 36(2), 115--132.

Kurlantzick, J. (2021). China's Myanmar Dilemma: Maintaining Stability and Assets Under the Junta. Council on Foreign Relations: Asia Program Brief, 14(2), 45-52.

Leivon, V. L. (2025, February 7). 2021 military coup in Myanmar and its impact along the India-Myanmar border region. Indian Council of World Affairs (ICWA).

Ministry of Commerce. (n.d.). About Myanmar. The Republic of the Union of Myanmar. <http://www.commerce.gov.mm>

Molthof, M. (2022). ASEAN and the principle of non-interference. Edinburgh University Press / International Relations of Southeast Asia.

Samtani, S. (2018, October 18). Deporting Rohingya refugees: Indian Supreme Court violates principle of non-refoulement. Oxford Human Rights Hub. <https://ohrh.law.ox.ac.uk/deporting-rohingya-refugees-india-supreme-court-violates-principle-of-non-refoulement/>

Singh, S. (2024). Bridging Boundaries: India's Neighborhood Policy Amidst Internal and External Challenge. New Delhi: Isara Solutions.

Sun, Y. (2024, January 16). Operation 1027: Changing the tides of the Myanmar civil war? Brookings Institute. <https://www.brookings.edu/articles/operation-1027-changing-the-tides-the-myanmar-civil-war/>

Tucker, S. (2023, July 25). Myanmar reveals ASEAN's weak spot again. Stimson Center. https://www.stimson.org

Wai Yan Phyo Naing, & Lin Sae-phoo (2025, may 09).Northern Myanmar Poses a Challenge to China's Critical Minerals Strategy. The Diplomat. <https://thediplomat.com/2025/05/northern-myanmar-poses-a-challenge-to-chinas-critical-minerals-stratergy/>

FIGURE 1

Areas under control of Junta and Local

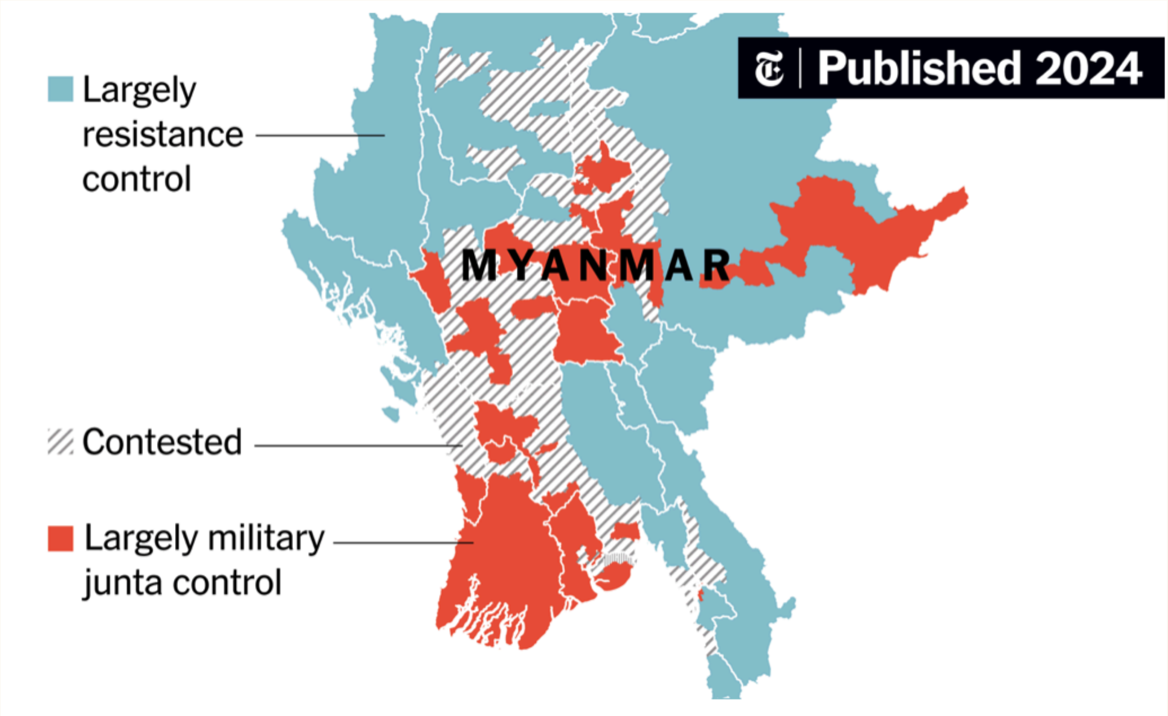


FIGURE 2

Major Projects made by China in Myanmar

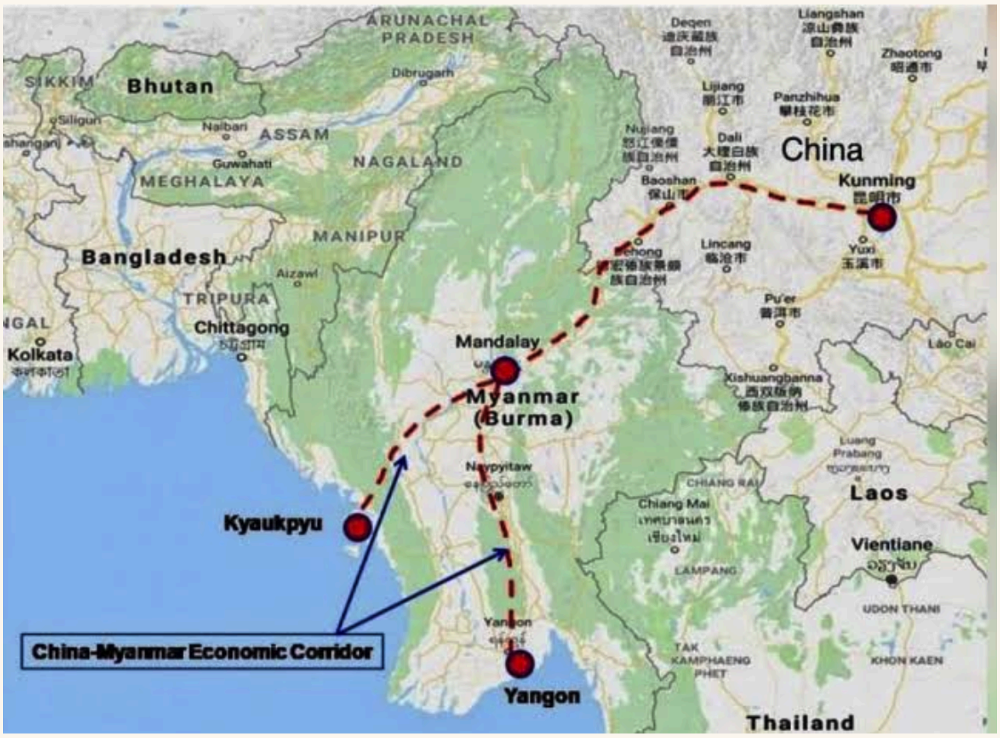


FIGURE 3

Kaladan Project in Myanmar by India

COLOPHON

This inaugural issue of The Exodus was set in Georgia and Arial and published on 14 July 2026 — Bastille Day — as Volume 1, Issue 1.

All papers are the work of high-school student researchers and represent the views of their authors alone.